Transmission of Respiratory Syncytial Virus in Households: Who Acquires Infection From Whom?

Thesis

Transmission of Respiratory Syncytial Virus in Households: Who Acquires Infection From Whom?

by

Patrick Kiio Munywoki

BSc, MSc

Affiliated Research Centre

KEMRI-Wellcome Trust Research Programme

Kilifi, Kenya

In collaboration with

University of Warwick, UK

DATE OF SUBMISSION: 28th June 2013

DATE OF AWARD: 30 SEPTEMBER 2013

Thesis submitted for the Degree of Doctor of Philosophy

The Open University, UK

Dedicated to

my adorable girls; Anne, Mercyln and Colette

Abstract

Households represent a setting of frequent and intense contacts and hence are conducive to the spread of respiratory viruses, such as respiratory syncytial virus (RSV). Infants are most vulnerable to severe RSV disease but a vaccine is not yet available hence the need to explore alternate strategies of protecting them. Such strategies would require better understanding of who infects the infants. During the RSV season of 2009/2010, we undertook a prospective study in rural Kenya involving 493 members of 47 households each with a child born after the preceding RSV epidemic and at least one elder sibling. Throughout the epidemic a nasopharyngeal swab (NPS) was collected every 3–4 days irrespective of symptoms, from all household members, and tested for a range of respiratory viruses including RSV using a molecular diagnostic assay. Partial sequencing of the attachment protein (G) gene from positive swabs was used to compare RSV strains within the household. In addition, once-a-week a specimen of oral fluid (OF) from around the gums was collected for RSV-specific antibodies screening and for assessment of sensitivity of the OF in detection of RSV using molecular diagnostics. This is the first prospective study to investigate introduction and transmission of RSV in families using molecular techniques over a complete RSV season. Analysis of RSV infection data is reported in this thesis with particular interest to identifying from where infants derive their infection, estimating the duration of RSV shedding and identify factors influencing the recovery rates, and estimating parameters of RSV susceptibility and transmission probability. In addition, data on diagnostic performance of OF in detection of RSV by molecular methods is presented.

A total of 16,924 NPS were collected, representing 86% of planned. RSV was detected in 40 (85%) households and 179 (36%) of the participants. In 28 of the 44 households with complete data, there was transmission of RSV to the infants experiencing their first epidemic. The probable source of RSV infection of the naïve infants was a household member in at least 54% of the cases. Co-primary infection between a household member and the RSV-naïve infant was ascribed in 4 of the cases. Older children were assigned the primary case for 11 (39%) of the infant cases and 10 (91%) of these were attending school. The infants appeared to play a role transmitting the introduced infections to the other members of the household including to the mothers.

These findings support vaccination strategies that target school age children and pregnant women. Both of these vaccination strategies can have profound benefits to RSV naïve infants directly by augmenting neutralizing antibodies against RSV (immunization of the pregnant women) and indirectly by reducing transmission from siblings to RSV-naïve infants. Results from this study provide increased confidence in the rationale for RSV vaccination of individuals who are not the key targets for protection.

Acknowledgements

I would like to first and foremost thank my supervisors, Professors James Nokes and Graham Medley for their guidance, constant support and encouragement throughout the project. Your mentorship has been invaluable and has helped me appreciate critical analysis and attention to detail in whatever I do. I have greatly learnt a lot and equally enjoyed working with you. It was a blessing to have you as my mentors.

Second, I would like to thank the very hard working 'household' study team: the study coordinator, Dorothy; the study clinician, Nuru; the field workers, Ben, John, Alfred, Patrick, Irine, Victor, Grace, Alan, Kalume, Edward, Christine; the laboratory staff, Clayton, Charles A, Carol, Ann, Alex, Regina, Clemet, James, Gertrude and Martin; and the data entry group, Shadrack, Eunice, and Grace, all of whom worked diligently with passion in their respective roles. Without your commitment and dedication to work, it would have been impossible to successfully complete the project. To the Matsangoni dispensary staff, KEMRI-community representatives, community health workers, chief and assistant chiefs, and other community leaders in involved in the study, I say thank you for your support in community engagement activities and giving a continuous feedback throughout the project. Special thanks goes to all the study participants for your voluntary commitment to the study. I appreciate the many times you welcomed my team and me to your households.

I would also like to thank the KEMRI- Wellcome-Trust Research Programme for granting me this opportunity and providing the necessary infrastructure to carry out the project. Special thanks to the 'Community Advice for Specific Studies' team, Vicki, Dorcas, Salim, Hassan and Isaac, from the Community Liaison Department.

Your guidance on ethics and community engagement was fundamental to the successful implementation of the project.

To my friends; Johnstone, Eric, Tim, Charles A, Charles S, Joyce, Patience, and Daniel, thank you for always being available to listen to me and for your constructive criticism. It was reassuring to always know I have you on my side.

Special thanks to my parents, Josephine and Munywoki, for from you I have inherited the sprit of hard work and discipline. Thanks to my brothers (James, Cyrus, Martin, Thomas and Dominic) and sisters (Catherine, Regina, Angeline and Winfred) who constantly kept me lively: let the spirit go on – and together we will prosper. Most of all, I would like to thank my wonderful wife, Anne, for her prayers and encouragement in every day of the four years and above all for her endless love. Last, but not the least, I would like to appreciate my magnificent daughters, Mercelline for cheering me up with tireless play in the house, and Colette, for conveniently 'delaying' her arrival until I returned from a study trip – I love you both.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS/ ACRONYMS

ARI	Acute respiratory infection
Ad or Adeno	Adenoviruses
CAST	Community Advice for Specific Studies
CCC	Communication and consent committee
CFR	Case fatality rate
CGMRC	Centre for Geographic Medical Research – Coast
CHW	Community health workers
CLG	Community Liaison Group
CI	Confidence interval
Ct	Cycle threshold
ELISA	Enzyme-linked immunosorbent assay
FI-RSV	Formalin-inactivated respiratory syncytial virus vaccine
HcoV	Human coronaviruses
HH	Household
HIV	Human immunodeficiency virus
hMPV	Human metapneumoviruses
HP	High performance, MagnaPure LC RNA isolation kit
IgG	Immunoglobulin G
IQR	Interquartile range
IFAT	Immunofluorescent antibody test
KDH	Kilifi district hospital
KEMRI	Kenya Medical Research Institute
KHDSS	Kilifi Health and Demographic Surveillance System
KWTRP	KEMRI-Wellcome Trust Research Programme
LAV	Live attenuated virus vaccine
LRTI	Lower respiratory tract infections

M-PCR	Multiplex polymerase chain reaction
MUAC	Mid-upper arm circumference
NERC	National Ethical Review Committee
NPA	Nasopharyngeal aspirate
NPS	Nasopharyngeal flocked swab
NW	Nasal wash
OF	Oral fluid
PERCH	Pneumonia Etiology Research on Child Health
PIV 1,2, 3, 4	Parainfluenza viruses type 1,2, 3 and 4
pfu	Plaugue-forming units
RNA	Ribonucleic acid
RSV	Respiratory syncytial virus
RV	Rhinoviruses
RT- PCR	Reverse transcriptase - polymerase chain reaction
SCC	Scientific Coordinating Committee
SSC	Scientific Steering Committee
TCID	Tissue culture infective dose
TNA	Total nucleic acid, MagnaPure LC total nucleic acid isolation kit
UK	United Kingdom
URTI	Upper respiratory tract infections
USA	United States of America
VEC	Viral Epidemiology and Control
WAIFW	Who acquires infection from whom
WHO	World Health Organization

CHAPTER ONE

1 General Introduction

1.1 Background of respiratory syncytial virus

Respiratory syncytial virus (RSV) is the major viral cause of childhood acute respiratory infection (ARI) worldwide. It is estimated about 60% of all children are infected during their first year of life and almost all by the end of their third year (Glezen *et al.* 1986). A recent review reported more than 30 million cases of RSV-associated ARI and at least 66,000 deaths occurred worldwide (99% of these deaths occurring in developing countries) in children under 5 years of age in 2005 (Nair *et al.* 2010). Studies undertaken in our setting of a rural developing country population typical of sub-Saharan Africa suggest around 1-2 in every 100 infants are admitted to hospital with RSV-associated severe pneumonia each year and RSV accounts for up to one fifth of all infant severe pneumonia, comparable to any other single respiratory pathogen (Nokes *et al.* 2009; Berkley *et al.* 2010).

Currently no RSV vaccines are licensed for use (Nokes and Cane 2008). Immunological immaturity and maternal antibodies represent major obstacles to vaccines targeting the key 1-3 month age group (Karron *et al.* 2005). Live attenuated vaccines in age groups 6 months and over are safe and immunogenic (Karron *et al.* 2005; Wright *et al.* 2007). However, the delay in delivery is at the cost of failing to directly protect a proportion of those most at risk. Hence, there is a need to carefully evaluate alternative approaches (Nokes and Cane 2008; Anderson *et al.* 2013), such as targeting any one or more of older infants, household siblings, school children or mothers. The potential impact of such alternative strategies will be intimately linked to who is acquiring infection from whom (WAIFW) (Anderson and May

1991), and in particular which groups are most important in infecting the infant. This forms the basis of the current study.

Household represents an important unit for the spread of infectious disease, and has been a focus for observational studies of respiratory illnesses in the past (Longini *et al.* 1982; Viboud *et al.* 2004). The close interactions in the household offer favorable environment for the spread of respiratory viruses. The household also represents an important setting through which demographic changes can translate to population effects on infectious disease transmission, for the spread of emergent infectious diseases, and on which to focus control strategies (Becker and Dietz 1995; Hall and Becker 1996; House and Keeling 2008; Goldstein *et al.* 2009; House and Keeling 2009). The primary interest in this thesis thus rests with understanding the transmission patterns of RSV in households. Given that disease following primary infection in infants is the major focus for vaccination, studies in the household setting are important in determining from where infants derive their infection and how these infections can be prevented.

The advent of broad-range molecular diagnostics provides sensitive tools for the study of the viral population dynamics which underpin disease aetiology (Gunson *et al.* 2005; Levine *et al.* 2012). Elucidation of within household transmission patterns of RSV calls for intensive sampling irrespective of symptoms using techniques of high sensitivity with particular demands of study infrastructure and community engagement, possible in few settings. In our settings, a nasopharyngeal specimen collected using the flocked swab has been shown to be acceptable in the community (Munywoki *et al.* 2011). However, oral fluid offers an alternative specimen for virus detection using either enzyme linked immunosorbent assay (ELISA) (Okiro *et al.* 2008) or molecular diagnostics (von Linstow *et al.* 2006). Data on the

diagnostic performance of this painless sampling method is limited and will thus be explored in the current study.

1.2 Objectives

This research feeds into a larger project whose central interest is to develop strategies for optimal vaccine intervention for the prevention of RSV disease. The main objective for this work is to define who acquires infection from whom within households in a developing country setting, with primary interest in the introduction and spread of RSV. The principle questions are related to who infects the infant of the household with RSV and which factors influence the duration of RSV shedding.

1.2.1 Primary objectives

a) Determine the proportion of primary infant infections that arise from within the household (e.g. from elder siblings, the mother, father and other household members) or from outside the household.

b) Estimate the duration of RSV shedding and identify factors influencing the rates of RSV recovery

c) Estimate transmission probability and susceptibility parameters key in defining WAIFW in the households.

1.2.2 Secondary objectives

a) Evaluate the use of oral fluid (OF) in detecting infection using multiplex PCR and ELISA methods

b) Determine the risk of subclinical RSV infections by age

1.3 Approach

During an RSV epidemic, which spanned December 2009 to June 2010, we undertook a study to investigate the introduction and spread of respiratory viruses in 50 households selected from a well-defined rural developing country community (Scott *et al.* 2012). Eligible households had an infant born since the previous RSV epidemic and one or more older siblings aged below 13 years. All members were swabbed twice weekly from the nasopharynx, and tested for a range of respiratory viruses using an in-house multiplex molecular diagnostic assay, and once a week from around the gums and assayed for RSV-specific antibodies. The sampling frequency set for this study was thought out carefully and arrived at after consideration of the different sampling intervals vis-à-vis shedding duration of RSV as outlined in Chapter 3 on study methods. The reasons for the broad virus screen on nasal swabs were: (i) identification of other respiratory viruses, which spread by a similar route to RSV, which would increase power of study to identify who infects whom, (ii) intrinsic interest in other viruses as causes of morbidity in the community, (iii) in order to assess the interaction of RSV and other respiratory viruses. Selection of the 'other' viruses to screen was based on the observation of the most prevalent during the study period. However, for the purpose of this thesis focus was on detection of RSV virus. With the RSV infection data, we quantify who acquires infection from whom and in particular identify from where the infants derive their infection. Elucidating the potential of different vaccination strategies and the impact at the population level of such approaches on RSV infection especially on the vulnerable infants will benefit from understanding of transmission dynamics of RSV.

1.4 Declaration of the author's role

The author was the principal investigator of the parent study (whose proposal is attached, appendix A) from which the analyzed data arise. He took lead role in conceptualization, design and implementation of the household study. The author addressed all scientific and

ethical issues raised by the research review boards in Kenya and UK. In terms of the study conduct, the author was the overall study manager. He was in charge of the field team, which consisted of one study coordinator, one study clinician, one senior field worker and 10 field workers; and data entry team, comprising of three data entry clerks. The author worked closely with the Laboratory Manager for the Viral Epidemiology and Control research group and the laboratory team in processing and screening as well as in storing of the study specimens. The custom-made Filemaker (version 11) database was developed and administered by the author. The study data were double entered and the author merged and cleaned the database. The author led all the analyses unless otherwise stated. The study was funded by Wellcome Trust grant (Grant number 090853).

During the period of registration, the candidate published the following articles:

a) Berkley JA, **Munywoki P**, Ngama M, Kazungu S, Abwao J, Bett A, Lassauniere R, Kresfelder T, Cane P, Venter M, Scott J, Nokes D. (2010) Viral etiology of severe pneumonia among Kenyan infants and children. JAMA 303: 2051-2057.

b) Nair H, Nokes DJ, Gessner BD, Dherani M, Madhi SA, Singleton R, O'Brien K, Roca A, Wright P, Bruce N, Chandran A, Theodoratou E, Sutanto A, Sedyaningsih E, Ngama M, **Munywoki P**, Kartasasmita C, Simoes E, Rudan I, Weber M, Campbell H. (2010) Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. Lancet 375: 1545-1555.

c) **Munywoki PK**, Hamid F, Mutunga M, Welch S, Cane P, Nokes D. (2011) Improved detection of respiratory viruses in pediatric outpatients with acute respiratory illness by real-time PCR using nasopharyngeal flocked swabs. J Clin Microbiol 49: 3365-3367.

d) Onyango CO, Welch SR, **Munywoki PK**, Agoti CN, Bett A, Ngama M, Myers R, Cane P, Nokes D. (2012) Molecular epidemiology of human rhinovirus infections in Kilifi, coastal Kenya. J Med Virol 84: 823-831.

e) Webb C, Ngama M, Ngatia A, Shebbe M, Morpeth S, Mwarumba S, Bett A, Nokes D, Seale A, Kazungu S, **Munywoki P**, Hammitt L, Scott J, Berkley J. (2012) Treatment Failure among Kenyan Children with Severe Pneumonia - a Cohort Study. Pediatr Infect Dis J 13: 13

f) Kamuya DM, Theobald SJ, **Munywoki PK**, Koech D, Geissler WP, Molyneux S. (2013) Evolving friendships and shifting ethical dilemmas: fieldworkers' experiences in a short term community based study in Kenya. Dev World Bioeth 13: 1-9.

g) **Munywoki PK**, Ohuma EO, Ngama M, Bauni E, Scott JA, Nokes D. (2013) Severe lower respiratory tract infection in early infancy and pneumonia hospitalizations among children, Kenya. Emerg Infect Dis 19: 223-229

h) **Munywoki PK**, Koech D, Agoti CN, Lewa C, Cane PA, Medley, GF, Nokes DJ; The source of respiratory syncytial virus infection in infants: a household cohort study in rural Kenya, submitted

i) **Munywoki PK**, Koech D, Agoti CN, Cane PA, Medley, GF, Nokes DJ; Duration of RSV shedding and factors influencing the recovery rates, in preparation

1.5 Overview of the thesis

Following this introduction is a comprehensive review of literature. The review covers RSV epidemiology and transmission with focus on providing background for the questions addressed in this thesis. A summary of the studies reporting spread of RSV in families is also presented, highlighting the key methodological considerations for the current study. The third chapter provides the detailed description of the study design and implementation and

challenges encountered. General results on household recruitment and retention, details on home visits and sample collections and prevalence of RSV in the inpatient surveillance over the study period are also presented. Chapter 4 provides further general results on baseline characteristics of the study households and individuals, seasonality of RSV and other respiratory viruses, probability of subclinical RSV infections by age and diagnostic performance of oral fluid in RSV detection using molecular techniques. Chapter five presents data on the source of RSV infection in study infants with particular focus on who introduces RSV into the households. Characteristics of infected and uninfected households and individuals are also presented and discussed. The sixth chapter describes the analysis of RSV shedding durations and factors influencing the infection recovery. In chapter seven using a mathematical model, transmission and susceptibility parameters are estimated. Results on the relative contribution by different household members on RSV infections are shown i.e. who infects whom. The final chapter provides a summary and discussion of the main findings stating the implications on strategies for vaccine delivery. The final chapter also discusses the study limitations while offering possible areas of improvement as well as directions for future research.

CHAPTER TWO

2 Literature Review

2.1 Summary

Respiratory tract infections are common and a major source of morbidity and mortality globally, particularly in low-income countries. About 155 million episodes of pneumonia occur in children younger than 5 years leading to the death of 1.5 million young children every year (Liu *et al.* 2012; Rudan *et al.* 2012; UNICEF *et al.* 2012). A recent review of 89 published and unpublished studies on acute lower respiratory tract infection (LRTI), estimated about 12 million hospitalisations for severe LRTI and 3 million for very severe LRTI occurred in 2010 worldwide in under five year olds. The majority (96%) of the episodes were in low-income countries. It is in this regard that childhood pneumonia has been identified as a priority for research (Adegbola 2012).

Data from aetiology studies identify *Streptococcus pneumoniae* and *Haemophilus influenzae* type b as the main bacterial causes of childhood pneumonia. Viruses are detected in up to 80% of pneumonia and bronchiolitis episodes in young children particularly in low-income countries (Gilani *et al.* 2012). Respiratory syncytial virus (RSV) and rhinoviruses are the most frequently isolated viruses (Berkley *et al.* 2010; Gilani *et al.* 2012) and RSV in particular has been identified as the main viral cause of infant and early childhood pneumonia and bronchiolitis worldwide. With improved access to childhood vaccines, such as *Haemophilus influenzae* type b, pertussis and pneumococcal conjugate vaccines, the role of viruses and in particular RSV will assume greater importance in efforts to reduce further the burden of childhood pneumonia (English and Scott 2008). An on-going multicentre study by the Pneumonia Etiology Research for Child Health (PERCH) group in seven low-income

countries will provide updated data on role of RSV in causing severe LRTI (Levine *et al.* 2012). The current study, will also add our knowledge on the role of respiratory viruses in causing acute respiratory infections (ARI) by identifying the circulating viruses in the community among potentially health individuals.

In a recent systematic review and meta-analysis, it was estimated 34 (19 – 66) million cases of RSV-associated ARI and 66000- 199000 deaths (99% of these deaths occurring in developing countries) occurred worldwide in children under 5 years in 2005 (Nair *et al.* 2010). The Global Burden of Disease study reported 253,500 (215000 – 296600) deaths due to RSV in all age groups in 2010 (Lozano *et al.* 2012). These estimates were higher than those reported by Nair e*t al* (Nair *et al.* 2010) probably due to lack of good quality mortality data and differences in study methods.

Disease due to RSV is principally the result of primary infection, which, because of high virus transmissibility, occurs predominantly in infants, and particularly in those under 6 months of age (Glezen *et al.* 1986; Weber *et al.* 1998; Nokes *et al.* 2004; Robertson *et al.* 2004; Nokes *et al.* 2008; Hall *et al.* 2009). Surveillance of paediatric admissions in rural coastal Kenya, Kilifi district, over the period 2002-07 has defined incidence estimates (per 100,000 per year) of RSV-associated severe or very severe pneumonia admissions of 1,107 (95% CI, 1012-1211) in infants, and 293 (271-317) in the under 5 year olds, contributing some 15% and 12%, by age group, respectively, of all admissions for these conditions (Nokes *et al.* 2009). This is comparable with the proportion of clinical severe pneumonia prevented by 9-valent conjugate pneumococcal vaccine in The Gambia (12%) (Cutts *et al.* 2005). Similar estimates of incidence of severe RSV disease have been observed in developed countries. In every 100 infants, about two experienced severe RSV disease requiring hospitalisation in their first year of life in United States and England (Sims *et al.* 1976;

Glezen *et al.* 1986). Within a birth cohort in the Kilifi Health and Demographic Surveillance System (KHDSS), followed up between 2002 and 2005, the incidences of RSV-associated pneumonia, severe pneumonia, and hospital admission among infants were 10400, 6600 and 1300 cases per 100,000 child years of observation, respectively. The proportion of cases of all-cause pneumonia, severe pneumonia and hospitalisations attributable to RSV in the cohort was 13%, 19% and 5%, respectively (Nokes *et al.* 2008). Hence, there is a considerable burden of RSV in the community not identified by hospital surveillance. Incidence of hospitalisation with RSV infection appears to decline with distance from the hospital, as demonstrated in studies from The Gambia (Weber *et al.* 2002) and also our inpatient studies (see Figure 2.1), which is most probably attributable to ease of access and health seeking behaviour. There is evidence that RSV predisposes to future respiratory disease in the short and longer term (Weber *et al.* 1999; Poulsen *et al.* 2006; Munywoki *et al.* 2013), adding to the ill-health burden. In summary, it is therefore apparent that there is a significant burden of RSV disease which warrants development of novel treatment and prevention strategies.

Glezen et al, in a birth cohort study, showed that approximately 60% of all children experienced RSV infection in their first year of life and almost all by their third year (Glezen *et al.* 1986). Furthermore, RSV repeatedly reinfects throughout life (Hall *et al.* 1976; Henderson *et al.* 1979). Undoubtedly this translates to a large pool of infectious individuals likely to be fundamental to endemic maintenance of infection and transmission to young children. The contribution of re-infections to the spread of RSV within the community is not well elucidated, but will be dependent upon the prevalence of individuals with re-infections, and their infectivity and patterns of contact with others in the population. These factors combine to define 'who acquires infection from whom' (WAIFW), which is central to

understanding the transmission dynamics of infectious diseases, and predicting vaccine impact (Anderson and May 1991).

No licensed RSV vaccine exists but promising live-attenuated vaccines are under development (Karron *et al.* 2005). With most of the severe RSV disease occurring in the first year of life, and particularly in the first 6 months of life, it is a priority that a vaccine protects this early age group. Live attenuated vaccines have been shown to be highly immunogenic and well tolerated in children 6 months and older, but poorly tolerated in children under 3 months of age without over-attenuation (Karron *et al.* 2005; Wright *et al.* 2007). It is apparent that alternative delivery strategies need to be explored (Anderson and Grenfell 1985; Collins and Murphy 2007; Anderson *et al.* 2013; Kaaijk *et al.* 2013), including delaying vaccination to older infants (>3 months) or targeting elder siblings, school children or mothers. The potential of such strategies to protect against severe RSV disease depends not only on the proportion of disease directly preventable in those receiving the vaccine but also on resultant reduced spread of infection from those vaccinated to those too young to receive vaccine (<3 months). Our studies identify over 60% of RSV severe disease occurring in children ≥3 months of age (Nokes *et al.* 2009); hence direct protection from delayed delivery may be substantial. Protection from infection afforded indirectly may be in the form of a reduced risk of transmission between presumed key transmission contacts (for example, mother to infant), or more generally due to reduced circulation of the infectious agent in the population, the so called herd immunity effect (Anderson and May 1991). These indirect protective and herd immunity effects are not quantified for RSV but well quantified for measles, influenza and hepatitis. Mathematical modelling of infection transmission has been previously used to evaluate this impact; exploring different vaccination strategies (Anderson and Grenfell 1986; Wilson *et al.* 2007; Vynnycky *et al.* 2008). However, such models require

well-determined parameters for WAIFW and particularly who is infecting vulnerable infants in the case of RSV. This study aims to delineate transmission patterns of RSV in the households with particular interest in identifying where infants derive their infection from (Chapter 5), investigating factors affecting RSV shedding durations (chapter 6), and quantifying parameters that define WAIFW matrix (Chapter 7).

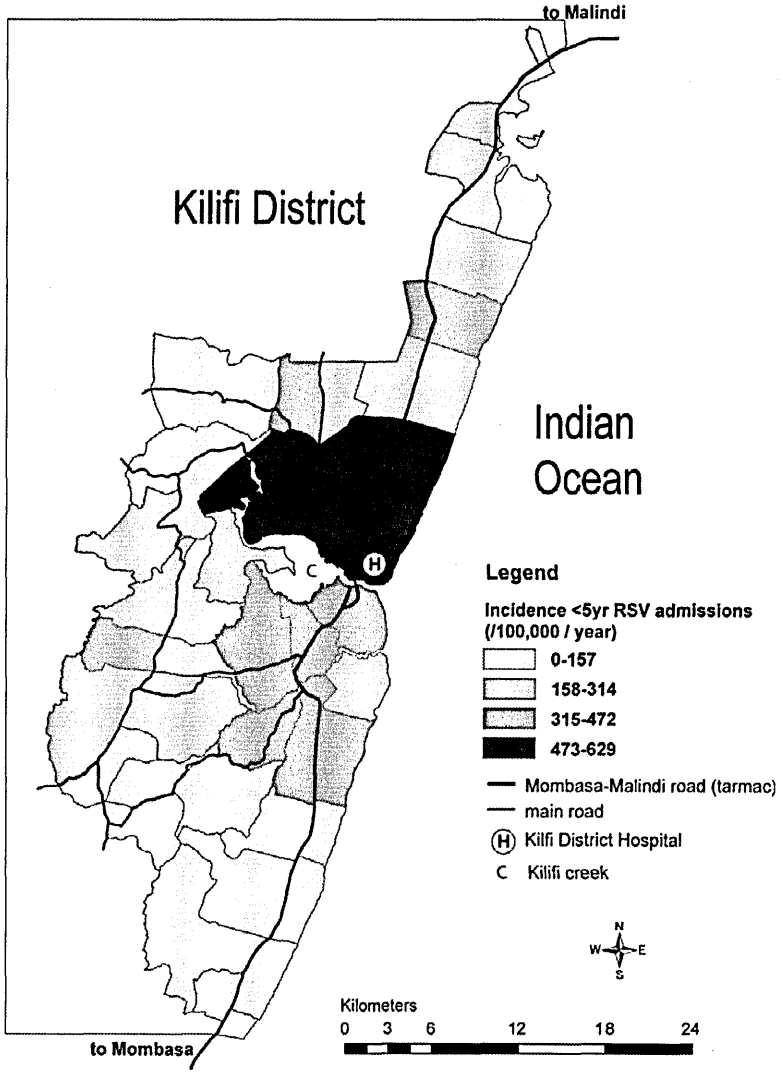


Figure 2.1: KHDSS map showing the incidence of RSV admissions to paediatric wards of Kilifi District Hospital (KDH) by administrative location. Source: Nokes et al, 2009 (Nokes et al. 2009)

2.2 Epidemiology of RSV

2.2.1 The virus

Human RSV was first isolated in children with respiratory illness in 1957 (Chanock *et al.* 1957), a year after being discovered in chimpanzees with coryza (Morris *et al.* 1956). The virus has since been antigenically characterised into two broad groups, A and B, initially by neutralisation tests using polyclonal sera (Coates *et al.* 1966) and monoclonal antibodies (Anderson *et al.* 1985) and later by gene sequencing (Collins *et al.* 1986; Johnson *et al.* 1987; Johnson and Collins 1988). RSV is an enveloped virus classified within the family *Paramyxoviridae*, genus *Pneumovirus*. The RSV genome is a non-segmented negative sense, single-stranded RNA molecule of approximately 15,200 nucleotides, which has 10 genes encoding for 11 distinct proteins. The proteins include two non-structural proteins, NS1 and NS2, which have been shown to antagonize interferon mediated antiviral responses (Cane and Pringle 1995). The nucleocapsid proteins are nucleoprotein (N) and phosphoprotein (P), essential for transcription, and the RNA polymerase (L). The virus has the matrix proteins M, M2-1 and M2-2 thought to be involved in viral assembly, transcription elongation and regulation of viral transcription, respectively. Three surface proteins are also encoded; small hydrophobic (SH) protein, glycoprotein (G) and Fusion (F) proteins. The role of the SH protein is not yet clear. The G protein is important in RSV attachment to cellular receptors (hence referred as the attachment protein) while F protein is the molecule responsible for the fusion of the virus and cell membranes and formation of a syncytium (Taylor *et al.* 1992). Both F and G proteins are major antigenic determinants and are involved in the acquisition of protective immunity against infection and disease (Piedra *et al.* 2003). The G protein has been shown to be more variable (about 47% variability in the amino acid sequences) compared to the F protein (20%) between the prototype RSV groups (Johnson *et al.* 1987).

There are several genotypes within each RSV group, distinguishable mainly by G gene nucleotide sequences, enhancing the antigenic variability of RSV to possibly allow re-infections to occur (discussed below). Studies investigating the genetic diversity of RSV have primarily focused on the variable G gene as reviewed by Cane (Cane 2007) and Melero (Melero 2007). In the current study, the long ectodomain region of the RSV G gene was sequenced to identify genetic relatedness of viruses circulating in the community and specifically within households.

2.2.2 Transmission characteristics

RSV is transmitted by close direct contact between individuals or material contaminated with nasal secretions (fomites), rather than over longer distances (e.g. two metres) by aerosol (Hall and Douglas 1981). Large droplets, not small, are required for spread presumably because it is an enveloped virus and labile to desiccation. Virus survival varies in relation to the surface or material type on which it is found and has been shown to survive for up to 6 hours on non-porous surfaces, such as countertops (Hall *et al.* 1980). Nosocomial spread of infections can occur when hospital staff touch secretions or contaminated objects while caring for an infected infant (Hall *et al.* 1975). With contaminated hands, self-inoculation may occur by the inadvertent rubbing of the nose or eyes, the major portals of RSV entry (Hall *et al.* 1981). The propensity for infection seems to depend on the amount of RSV inoculum transmitted as reported in an experimental challenge study (Lee *et al.* 2004) and, to a large extent, on the susceptibility of the host, measured by the amount of neutralising antibody titres prior to inoculation (see section on RSV immunity below) (Hall *et al.* 1981; Lee *et al.* 2004). The fact that RSV spreads via close contact makes settings such as households and schools important in studying RSV transmission and delineating RSV epidemiology.

2.2.3 *Seasonality of RSV*

RSV circulation occurs periodically and in most places every year albeit with considerable variation of the peak months of the RSV outbreak. Studies in Birmingham, UK, reported RSV outbreaks peaking in December every year (Cane *et al.* 1994) like in other parts of Europe (Mlinaric-Galinovic *et al.* 1994). In North America, RSV outbreaks appear to peak in February (Felton *et al.* 2004). The pattern is different in Finland with RSV activity occurring every two years with minor peaks in April followed by a major peak in December (Waris 1991). Recent studies in Croatia, have reported similar biennial cycle with one larger and one smaller RSV outbreak every 23-25 months (Mlinaric-Galinovic *et al.* 2008; Mlinaric-Galinovic *et al.* 2012). In the tropics most of the RSV epidemics occur annually peaking mainly between November and January (Reese and Marchette 1991; Bedoya *et al.* 1996; Chew *et al.* 1998; Weber *et al.* 1998; Chan and Goh 1999; Chan *et al.* 1999; Ahmed *et al.* 2012). The factors influencing the heterogeneity in timing of the RSV seasons are unclear. In temperate countries epidemics appear to occur in winter while in tropics the association with meteorological factors is less clear. Studies in countries in Asia, Africa and South America have reported association of RSV outbreaks and the wet season. There is contradicting evidence of RSV occurring in dry seasons (Cherian *et al.* 1990; Bedoya *et al.* 1996; Chan *et al.* 1999). In a retrospective study of samples positive for RSV in Singapore carried out from 1990 to 1994, the peaks were not associated with rainfall but with higher environmental temperature, lower relative humidity and higher maximal day-to-day temperature variation (Chew *et al.* 1998). A population-based surveillance of RSV in 1999 to 2001 among children younger than 5 years in Indonesia, South Africa, Nigeria and Mozambique using a standardised WHO protocol, reported peaking of RSV outbreak in rainy season in Indonesia and Mozambique while in Nigeria and South Africa RSV occurred primarily in the dry

season. The mechanism of the seasonal triggers of RSV in both temperate and tropical countries has been hypothesized to relate to changes in airway physiology, socio-behavioural patterns, and favourable conditions for the virus survival. In winter susceptible individuals crowd indoors and there is decreased protection in respiratory mucosa due to cooling of nasal passages. In the tropics, children tend to be kept indoors during rainy season, and the resultant crowding may also account for the increased transmission of RSV in this period. Others have suggested high humidity (in the wet season) may be favourable to viral survival by preventing drying and loss of infectivity of the virus (Hall and Douglas 1981). RSV is known to be a labile virus, and does not survive well under high temperatures which may explain the relationship with cooler weather. However, this trend was not observed in the Singapore study (Chew *et al.* 1998). Genetic variation of RSV viruses coupled with waning of RSV immunity and accumulation of susceptible individuals at the population level offer alternative explanations to the seasonal nature of RSV epidemics, as discussed in sections below.

2.2.4 Molecular epidemiology of RSV

The seasonal occurrence of RSV is associated with considerable genetic variation of the circulating strains. Evidence from most of the longitudinal studies show both of the RSV groups co-circulate during the RSV outbreaks although in most epidemics one group predominates. Studies in different regions have identified varying but consistent patterns of circulation of RSV group A and B with sequential replacement of genotypes. In the Birmingham, UK, a cyclic triennial pattern in groups A and B was observed from 1989 to 2000. Group A isolates were most commonly detected in eight epidemics with group B predominating every third year. Similar patterns have been observed elsewhere (Zlateva *et al.* 2007). In Finland, group dominance alternates every two years (Waris 1991). In general

group A genotypes have been found to predominate in most epidemics as has been observed in our settings (Figure 2.2). Multiples genotypes are detected in each epidemic with yearly replacement of the predominating genotype (Cane *et al.* 1999; Cane 2001; Roca *et al.* 2001; Seki *et al.* 2001; Zambon *et al.* 2001; Scott *et al.* 2006; Zlateva *et al.* 2007). Analysis of strain variability within epidemics has almost exclusively focused on isolates from hospitalised infants. However, studies comparing RSV variability in hospital with field isolates found similar viruses were circulating in the two populations (Zambon *et al.* 2001; Venter *et al.* 2002; Nokes *et al.* 2004; Scott *et al.* 2004). In South Africa, the same strain of RSV was found to cause both mild ARI and severe LRTI (Venter *et al.* 2002). These observations indicate inpatient surveillance of RSV could provide information on the epidemics in general. In our study, we used district hospital inpatient RSV surveillance to identify the start and end of the RSV epidemic in the community, thus enabling the timed focus of resources for active home follow up and sample collection to when the virus was circulating in the community.

As noted earlier the G gene is highly variable and is the target for most molecular epidemiology studies. The variations in the G gene present an opportunity to track RSV transmission and dissemination (Cane *et al.* 1991; Cane 2007). A review on molecular evolution of RSV by Cane showed new RSV genotypes are continuously emerging but also some previously circulated ones have disappeared (Cane and Pringle 1995). The extreme of RSV evolution was demonstrated in the occurrence of a RSV group B genotype (referred to as BA variant) with a 60-nucleotide duplication in the variable region of the G gene. The BA variant was first observed in South America in 1999 (Trento *et al.* 2003). The variant which probably arose from a single mutation event has increasingly become the dominant group B variant worldwide in recent years (Zlateva *et al.* 2005; Scott *et al.* 2006; Trento *et al.* 2006;

Zlateva *et al.* 2007; Trento *et al.* 2010; Agoti *et al.* 2012). It is thus apparent when new strains occur, they can rapidly spread around the world. Another review by Cane et al showed very similar viruses circulate worldwide at the same time (Cane *et al.* 1992) and viruses isolated in geographically distant places and in different years may be more closely related genetically than viruses isolated in the same location in the same week (Cane *et al.* 1999).

The mechanism for the change of dominant strains each year is not yet established but may be influenced by the build up of immunological resistance in the community to successive epidemics of the same strain i.e. accumulation of herd immunity. Emergence of new variants might also play a role in occurrence of epidemics. Social networks including travelling patterns could help in the global dissemination of the viruses. At a local scale, understanding the transmission patterns of RSV in households, which represents the smallest social unit in the community with intense interactions, would provide valuable insight into the RSV spread.

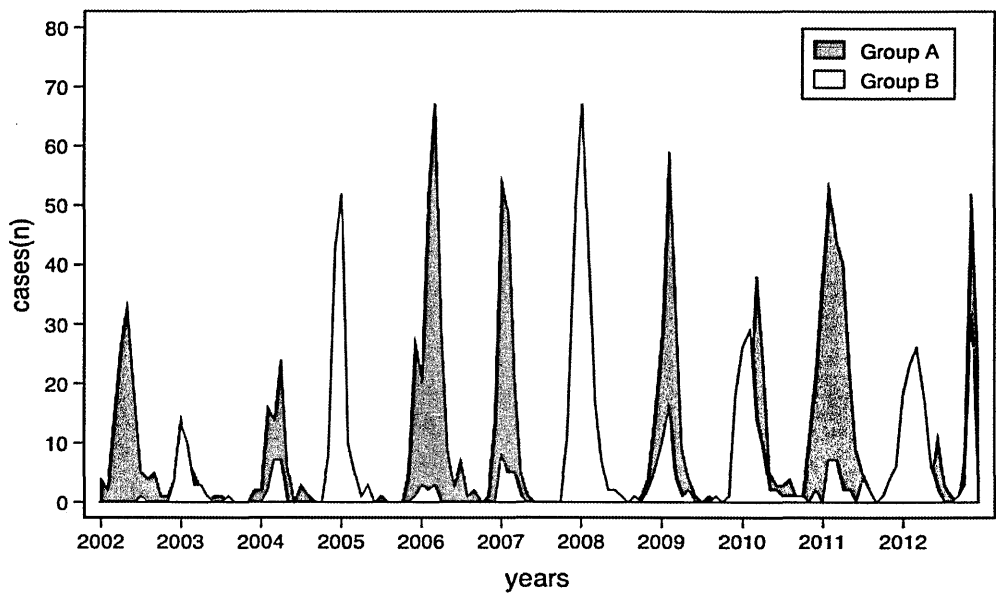


Figure 2.2: Seasonal variation of RSV groups A and B in children admitted to Kilifi District Hospital, Kenya, from January 2002 to December 2012

2.2.5 RSV disease burden

RSV is associated with a wide spectrum of respiratory symptoms ranging from mild upper respiratory tract illness (URTI) such as common cold to severe LRTI, mainly pneumonia and bronchiolitis in younger children. Published pneumonia aetiology studies have identified RSV as the main viral cause of pneumonia in children (Forgie *et al.* 1991; Berkley *et al.* 2010; Hammitt *et al.* 2012). In The Gambia, RSV was frequently detected among cases (49%) relative to the controls (19%) in a community study involving 90 cases of pneumonia in infants (<1 year old) and 43 age-matched community controls. The study involved a small number of cases in a restricted age group (<1 year) but other studies with bigger sample sizes have replicated these results as reviewed by Weber and colleagues (Weber *et al.* 1998). A pneumonia aetiology study in rural Kenya, detected RSV in 34% of the 760 severe and very severe pneumonia paediatric (<15 years) admissions and in 5 % of 56 healthy community controls using highly sensitive molecular techniques (Berkley *et al.* 2010). RSV was the only virus significantly associated with pneumonia admissions in children younger than 5 years (odds ratio, 12.5; 95% confidence interval, 3.1 – 51.5) in another recent study in the same Kenyan settings (Hammitt *et al.* 2012). The second Kenyan study employed a more rigorous study design and methods (Hammitt *et al.* 2012). The evidence taken together firmly implies that RSV is a major player in causing LRTI in children.

i) RSV infection and disease in children

RSV is a highly contagious virus with infections occurring in the early months after birth. Birth cohort studies show approximately 60% of newborns are infected before the first birthday with approximately 20% of these infections being associated with symptoms of LRTI. Almost all children are infected by their third birthday. This was demonstrated in a classic study in Houston, Texas, where 125 newborns were actively followed at home from

birth to the age of 5 years with collection of nasal samples when symptoms of ARI were recognized (Glezen *et al.* 1986). The nasal samples were tested for RSV using viral culture. In the Houston study, 69% of infants had RSV infection in the first year of life; of which a third developed LRTI. In the second year, 83% of children were infected and 16% were associated with LRTI. Even during the third and fourth years, a third to half of the children were infected, with LRTI occurring in a quarter of these infection (Glezen *et al.* 1986). Repeated infections were thus not uncommon and in some instances occurred within one year or less (Glezen *et al.* 1986). Another birth cohort in Kilifi, Kenya, reported a risk of ~40% and 10% of primary and secondary RSV infections in infancy (Nokes *et al.* 2004). The risks were lower than those in the Houston family study perhaps due to differences in study methods and population. The Houston study used virus detection in culture and 4-fold rise in Neutralising antibodies in paired sera while the Kilifi study used immunofluorence antibody test (IFAT).

The incidence of RSV infection and disease in children is largely similar in both developed and developing countries. Two birth cohorts, one in the low-income urban families in the Cali, Colombia, and the other in rural Kilifi, reported a rate of RSV associated LRTI of 224 and 154 per 1000 child-year (cy) in infants respectively (Borrero *et al.* 1990; Nokes *et al.* 2008). In the Kilifi birth cohort study the incidence of severe RSV disease in the community (~100/1000 cy) was approximately ten times that seen in the inpatient settings (~13/1000cy) suggesting that a considerable proportion of severe RSV cases within the community does not present in health facilities (Nokes *et al.* 2008). A finding that may be related to poor hospital access and health seeking behavior particularly in developing countries. The denominator-based studies reviewed in Table 2.1, show the incidence of RSV-LRTI to be consistently higher in infants relative to the older children (<5 years) in same populations

(Berman 1991; Weber *et al.* 2002; Djelantik *et al.* 2003; Robertson *et al.* 2004). There is considerable variation in reported incidences in different population settings ranging from 9/1000cy in Western Gambia to 220/1000 cy in Cali, Columbia in infants and from 3/1000cy in Hong Kong, China, to 116/1000cy in Ibadan, Nigeria among the under five year olds. Differences in study design (community versus hospital surveillance), location (rural vs. urban), year of study, specimen collection method (nasopharyngeal aspirate versus nasal wash or nasal swab), methods of RSV detection (culture versus IFAT or molecular techniques), surveillance methods (active versus passive), definition of LRTI (mild vs. severe or very severe) among others, play a role in this variation (Nokes 2007).

Since 2002, we have conducted a continuous surveillance of RSV in the Kilifi district hospital (KDH) paediatric wards (Nokes *et al.* 2009). Children admitted with clinical features of WHO defined severe or very severe pneumonia or clinician's diagnosis of LRTI are recruited and a nasal wash or NPS collected soon after admission. The samples are tested for RSV by IFAT and real time multiplex polymerase chain reaction (M-PCR) from 2007 distinguishing the RSV groups in circulation. Figure 2.3 shows the age distribution of the all the RSV-associated admissions from January 2002 to December 2012 showing an association of burden of severe RSV disease and age. Children aged less than 3 months experience the highest burden of severe RSV disease but an appreciable burden in the older age groups exists. Approximately 64% of severe RSV-LRTI has been in children aged 3 months or more and 43% in children aged 6 months or more (Nokes *et al.* 2009) (plus unpublished data).

Globally, 22% of all episodes of acute LRTI among the under five year olds are associated with RSV translating to about 34 million episodes of RSV-associated LRTI per year and 3 million episodes requiring hospital admission (Nair *et al.* 2010). The burden of RSV-LRTI in children is thus substantial and effective RSV vaccines could be the next 'frontier', after

introduction of pneumococcal conjugate vaccines, with great potential of reducing the LRTI morbidity and mortality, worldwide.

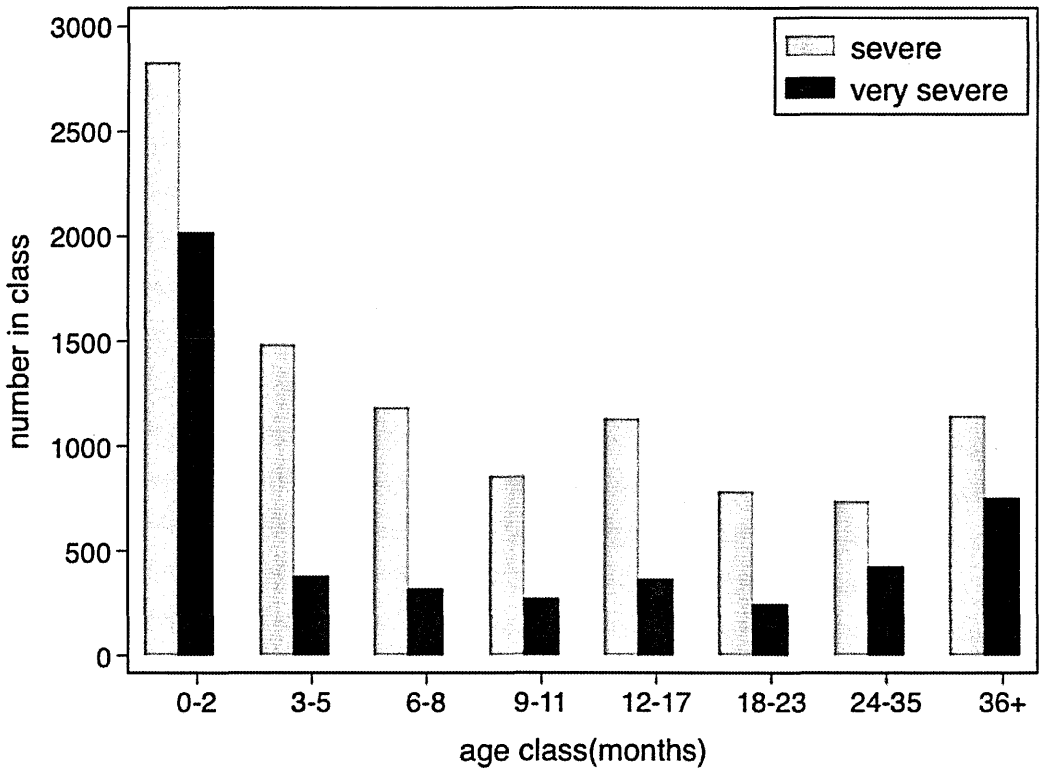


Figure 2.3: Age distribution of RSV associated severe and very severe LRTI admissions of children under 5 years old to KDH paediatric wards from January 2002 to December 2012

Table 2.1: Incidence estimates of RSV-associated LRTI per 1000 child years from studies in resource limited settings

Location				
i) Inpatient[1]	**Age class (yrs)**	**Cy**	**RSV incidence per 1000 cy (95% CI)**	**Reference**
Alaska, USA	<1	1,801	155 (137 – 173)	(Karron *et al.* 1999)
Negev, Israel	<1	2,991	18 (13 – 23)	(Dagan *et al.* 1993)
Lombok, Indonesia	<1	15,000	25 (22 – 28)	(Djelantik *et al.* 2003)
	<2	30,000	14 (13 – 15)	
Hong Kong, China	<5	496,000	3 (2 – 3)	(Chan *et al.* 1999)
Soweto, South Africa	<2	24,000	4 (3 – 4)	(Madhi *et al.* 2000)
Manhica, Mozambique	<1	1,342	15 (8 – 22)	(Robertson *et al.* 2004)
	<5	6,020	5 (3 – 7)	
Western Gambia (urban)	<1	20,338[4]	9	(Weber *et al.* 2002)
Western Gambia (rural)	<1	-	18	(Weber *et al.* 2002)
Kilifi, Kenya	<1	311	13 (0 – 25)	(Nokes *et al.* 2004)
ii) Outpatient[2]				

Cali, Colombia	<15	18,225	6 (5 – 7)	(Berman *et al.* 1983)
	<5	10,329	9 (7 – 11)	
	<1	2,066	25 (18 – 31)	
Manhica, Mozambique	<1	1,342	30 (21 – 39)	(Robertson *et al.* 2004)
Agincourt, South Africa	<1	1,652	15 (9 – 21)	(Robertson *et al.* 2004)
	<5	8,258	9 (7 – 11)	
iii) Community[3]				
Rio de Janeiro, Brazil	<5	786	14 (6 – 22)	(Sutmoller *et al.* 1995)
Cali, Colombia	<1.5	413	200 (149 – 241)	(Borrero *et al.* 1990)
	<1	399	220 (156 – 284)	
Ibadan, Nigeria	<5	1,579	116 (78 – 154)	(Robertson *et al.* 2004)
	<1	316	94 (79 – 109)	
Kilifi, Kenya	<1	311	154 (111 – 198)	(Nokes *et al.* 2004)
Takhli, Thailand	<5	14,569	8 (7 – 10)	(Suwanjutha *et al.* 2002)
Bandung, Indonesia	<1	284	41 (17 – 65)	(Robertson *et al.* 2004)

	<5	1,420	34 (24 – 44)	
Metro Manila, Phillipines	<5	1,418	28 (19 – 37)	(Tupasi *et al.* 1990)

Key: 1, RSV-LRTI diagnosed on admission to hospital; 2, LRTI identified through passive surveillance in out-patient department/clinic; 3, RSV-LRTI identified through active surveillance of the community; cy, child-years of observation; CI, confidence interval; Table adapted from a review by Nokes et al (Nokes 2007)

ii) RSV infections and disease in adults

RSV infection data in adults is limited with a few studies in developing countries (Karstaedt *et al.* 2009; Olsen *et al.* 2010; Feikin *et al.* 2012). In Western Kenya, the incidence for RSV-associated ARI was 0.98 per 100 person-years among HIV positive adults (>18 years) and 0.13 per 100 person-years in HIV negative adults. A study in South Africa did not find an excess of hospitalisation or mortality in adults during RSV seasons (Karstaedt *et al.* 2009). In developed countries, RSV infections are not uncommon in adults particularly in those with cardiac, pulmonary and immunodeficiency diseases and the elderly (Falsey and Walsh 2000). Outbreaks of RSV in long-term care facilities and studies in hospitalised adults have suggested RSV may be important cause of illness in community-dwelling elderly people (Zambon *et al.* 2001). A prospective cohort evaluated all respiratory illnesses using culture, serology and RT-PCR in healthy elderly patients and high-risk adults (those with chronic heart and lung disease) and inpatients hospitalized with acute cardiopulmonary conditions during 4 consecutive winters. The annual prevalence of RSV infections was 3 – 7% and 4 – 10% among the healthy elderly patients and the high-risk adults, respectively. Based on discharge diagnoses, RSV infection accounted for 11% of admissions with pneumonia, 11% with chronic obstructive pulmonary disease, 5% with congestive heart failure, and 7% with asthma. The authors estimated RSV infection would approximately account to 177,000 and 14,000 admissions and deaths annually, respectively, in the US. Thompson et al. have estimated that RSV accounts for approximately 10,000 deaths annually in the United States in persons over the age of 65 years using mathematical models linking viral activity in children with hospitalization and death in adults (Thompson *et al.* 2003). RSV related outpatient visits made by elderly adults for respiratory illnesses have been reported to be high during winter; up to 18% in the UK (Zambon *et al.* 2001). The burden of RSV disease in this age group cannot be disregarded as it is associated with considerable resource utilization. It is

estimated hospitalization costs would exceed $1billion and outpatient costs would be also substantial. RSV vaccines targeting this particular age group are thus warranted.

iii) RSV mortality

As highlighted earlier on, data is scarce on RSV associated mortality. The available data suggest mortality resulting from RSV is low with most studies reporting case fatality rates (CFR) of zero (range, 0 – 9%) (Weber *et al.* 1998; Stensballe *et al.* 2003; Nokes 2007). However, the mortality appears higher among children with underlying medical conditions (CFR of 3 – 5%) such as those with chronic lung disease, congenital heart disease and immune deficiency (Wang *et al.* 1995) and in developing countries (CFR of up to 9%) (Stensballe *et al.* 2003). Studies from developed countries report a CFR of less than 0.1%, reviewed in Stensballe et al (Stensballe *et al.* 2003). It is possible the reported RSV mortality is under-estimated in low income countries, since it has been reported that most LRTI associated deaths occur before a nasal specimen can be collected (Djelantik *et al.* 2003; Nokes *et al.* 2009) or outside the clinical setting at home where the cause of death may not be ascertained (Moisi *et al.* 2011).

2.3 Immunity to RSV infection and disease

RSV infection elicits both humoral and cellular immune responses. The role of cellular responses appears to be in enabling viral clearance during infection while antibodies response are associated with resistance to infection (Lee *et al.* 2004) and disease (Piedra *et al.* 2003). The Lee *et al* study showed a correlation between the level of pre-challenge neutralizing antibody titer and risk of a successful infection (Figure 2.4). Young adults with low neutralizing antibody titres had a higher risk of infection compared to those with high titers (Lee, Walsh et al. 2004). RSV infection induces antibodies against a variety of viral proteins but it is only those that are against the F and G proteins that have been shown to be neutralizing in vitro and protective in vivo (Taylor *et al.* 1984; Walsh *et al.* 1984; Connors *et*

al. 1991). Antibodies against the F protein are strongly cross-reactive against the heterologous RSV group (Hendry *et al.* 1988). However, those against the G protein show limited cross-reactivity between the two RSV groups (Hendry *et al.* 1988; Scott *et al.* 2007; Taylor 2007). The RSV G gene is characterized by the accumulation of amino acid changes at estimated rate of 0.25% per year for the whole molecule (Cane and Pringle 1995). In addition, the attachment protein has the highest ratio of nonsynonymous-to-synonymous nucleotide mutations (Cane and Pringle 1995) and number of positively selected sites of all RSV genes, suggesting these changes and sites might be immune driven (Botosso *et al.* 2009). A recent study in Kilifi, Kenya investigated the RSV group-specific neutralizing antibody responses to both contemporary and historical test viruses (Sande *et al.* 2013). The results showed that individuals had stronger neutralizing antibody responses to test viruses of the same infecting group (homologous seroconversion of 40 - 50%) than to viruses of the alternative group (heterologous seroconversion of 8 – 12.5%). The group-specific neutralizing response was not lost even after significant genetic changes within the RSV groups. The 60-nucleotide duplication (BA variant), which was thought to be a phenotype resistant to previous host group-specific immunity due to its increased global transmission (Trento *et al.* 2010), did not confer an ability to escape neutralizing responses to previous non-BA viruses *in vitro* (Sande *et al.* 2013). Together these data imply that the observed cyclic alternation in circulation of RSV group A and B could be due to population-level group-specific immunity and the effect of future RSV vaccines at reducing the infection would be enhanced if they contain representative strains from both group A and B.

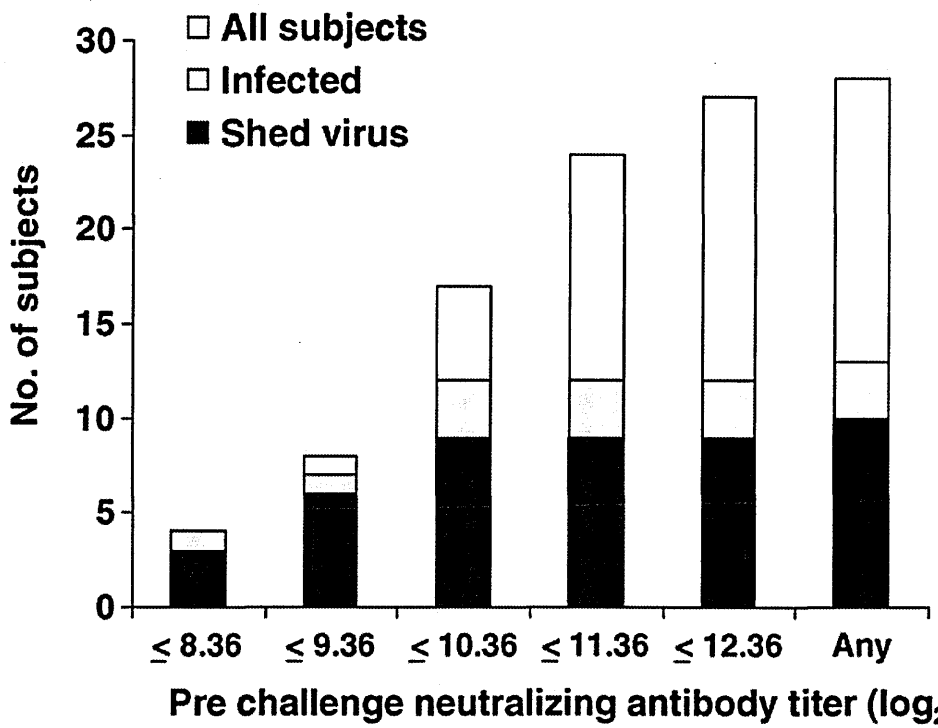


Figure 2.4: The relationship between pre-challenge RSV neutralizing antibody titres in serum and risk of infection. The graph shows the cumulative number of subjects at each neutralizing antibody titre who shed RSV or were infected. Infection detected by viral culture or a ≥four-fold response in serum antibody (adapted from Lee et al (Lee et al. 2004))

2.4 Subclinical RSV infections

Evidence from experimental challenge studies in adults and epidemiological studies demonstrate subclinical RSV infections are common in older populations. Results of an experimental study using RSV A2 strain in young adult volunteers, reported a greater risk of infection in individuals administered with a high dose of RSV inoculum compared to those with low dose (57% vs. 36%) with asymptomatic infections occurring in 63% and 40 % of the infections, respectively (Lee *et al.* 2004). The authors observed that infected individuals had statistically lower serum neutralising antibody titres (9.5 $\log_2$ vs. 10.6, p=0.04) than uninfected individuals (Table 2.2). In another experimental infection study involving 13 adult volunteers, with low neutralising antibody titres, 12 had detectable viral shedding and

respiratory illness when administered a high dose of RSV A2 inoculum (Lee *et al.* 2004). The prevalence of symptomatic infection is high in families exposed to RSV, which could be due to exposure to large inoculum from close social contacts within the household (Hall *et al.* 1976). In the Hall family study, only two of the 39 infected individuals had asymptomatic infection (Hall *et al.* 1976). Crowcroft *et al* reported a high prevalence of asymptomatic of 48% and 49% in siblings and parents, respectively, of infants admitted to intensive care units with RSV (Crowcroft *et al.* 1999). The Hall et al study (Hall *et al.* 1976) might have missed subclinical infection particular in older children and adults. In general, it appears subclinical RSV infections are likely in two scenarios: First, when individuals are exposed to low viral load of RSV such as in inter-family contacts as for the case of introducers of RSV into the households (Hall *et al.* 1976). Second, among individuals with high levels of neutralizing antibodies such as older children and adults (Lee *et al.* 2004). In order to identify chains of transmission in households it would be necessary to collect samples regardless of symptoms and test using very sensitive molecular based techniques to minimize the probability of missing infection episodes. The risk of asymptomatic infection, by age, is estimated in the present study. This is of interest as it has implications on the conduct and interpretation of RSV epidemiological studies. If individuals change their behaviour as a result of symptoms e.g. reducing their contacts, the period of shedding before symptoms appear or prevalence of asymptomatic infection is likely to have a role in transmission of RSV in the population.

Table 2.2. Adult challenge studies showing risk of infection and subclinical infection

Reference	Population	Inoculum	Infected[1] n (%)	Sub-clinical[2] n (%)	Comments
(Kravetz *et al.* 1961)	41 adult males	160 to 640 $TCID_{50}$ virus	33 (80.5)	13 (39.4)	
(Lee *et al.* 2004)	14 healthy young adults	4.7 log_{10} $TCID_{50}$ virus	8 (57.1)	5 (62.5)	
	14 healthy young adults	3.7 log_{10} $TCID_{50}$ virus	5 (35.7)	2 (40.0)	
	13 healthy young adults	4.7 log_{10} $TCID_{50}$ virus	12 (92.3)	0 (0)	The adults had low titres (<=9.36 log_2)
(Mills *et al.* 1971)	16 adult males	Low dose, 500pfu	16 (100)	16/16 (100)	
	17 adult males	High dose, 100000pfu	9 (52.9)	4/9 (44.4)	
(Hall *et al.* 1991)	15 adults	Repeatedly challenged after natural infection	15 (100)	15 (49% of the total challenge infections)	Natural challenge for up to 6 times

Key: 1, number (n) and percentage (%) of the infected out the total inoculated with the virus; 2, number (n) and percentage (%) of the asymptomatic infections

2.5 RSV reinfections

The Houston family study and other community studies report RSV episodes in older children and adults with history of previous infections (Berglund 1967; Monto and Cavallaro 1971; Cooney *et al.* 1975; Hall *et al.* 1976; Glezen *et al.* 1986; Nokes *et al.* 2008; Agoti *et al.* 2012). In the Houston study, 76% of the young children infected in the first year of life were reinfected in the second year and overall 53% (98/185) of the children had a reinfection in the follow up period (Glezen *et al.* 1986). Hall *et al* reported substantial attack rates in all age groups; 17% in adults compared to 29% in infants (Hall *et al.* 1976). Frequent and potential ease of repeat infections even with the same viral strain has been demonstrated in a challenge study in adults (Hall *et al.* 1991). In the study, over 70% of the adults were re-infected two times or more after a natural infection (Hall *et al.* 1991). Additional studies have shown RSV to be an important cause of respiratory tract infection in adults and especially in the elderly (Falsey and Walsh 2000). The mechanism for repeat infections is perhaps due to waning or partial immunity (Ohuma *et al.* 2012) and/or due to antigenic variation of the virus (Cane 2007; Agoti *et al.* 2012). Primary infections and, independently, young age have been associated with higher risk of severe disease relative to re-infections and are more likely to lead to hospital admission (Henderson *et al.* 1979; Glezen *et al.* 1986; Nokes *et al.* 2004; Nokes *et al.* 2008). However, re-infections are likely to constitute the bulk of RSV infections during the annual epidemics observed worldwide. It follows that re-infections probably play a major role in the overall transmission of RSV in the population. The importance of RSV infected older children and adults, many experiencing repeat infection, in the transmission of virus to infants may be considerable and needs to be quantified. Estimates are required of the incidence of infection in these age groups, the level and duration of viral shedding. Probability of reinfections occurring within the same epidemic and characterisation of the re-

infecting virus will be assessed in the current study. It would be important to assess whether there is any preponderance for reinfection with different genotypes.

2.6 RSV shedding

Current understanding on the patterns of RSV shedding following natural infections arise from studies in developed countries (Hall *et al.* 1975; Hall *et al.* 1975; Hall *et al.* 1976; Hall *et al.* 1976; Hall 1977; Frank *et al.* 1981; Waris *et al.* 1992; Hall *et al.* 2001) with only one exception (Okiro *et al.* 2010). Table 2.3 presents a summary of the studies reporting duration of RSV shedding. The data suggest durations of RSV shedding are of the order of days. However, the duration seems to vary depending on individual characteristics such as age and disease severity (Hall *et al.* 1976). A study involving 23 children hospitalized with respiratory illness (18 pneumonia, 5 URTI) followed until shedding ceased, reported an average duration of RSV shedding of 6.7 days (range 1 – 21 days) (Hall *et al.* 1976). In this study shedding duration did not correlate with age but with disease severity (Hall *et al.* 1976). Children with LRTI shed for a significantly longer period than those with URTI (mean duration of 8.4 days vs. 1.5 days). Other markers of disease severity such as pulmonary consolidation and the presence of underlying cardiopulmonary disease were also significantly associated with longer durations of shedding (mean 11.5 vs. 5.2 days and 9.9 vs. 5.4 days, respectively) (Hall *et al.* 1976). This was a hospital-based study with small samples sizes to be very sure of the differences observed and likely to provide biased estimates towards infections causing severe disease and/or of younger age groups. Inpatient studies in general suffer from additional limitations related to the observation that RSV shedding starts prior to clinical symptoms (Frank *et al.* 1981) and continues in most cases after clinical recovery (Hall *et al.* 1975; Hall *et al.* 1976). It is thus plausible the actual durations are longer for the hospitalized patients than those reported. Community studies offer best estimates. One such study following 36 families in Rochester, United States of America, through home visits by a

nurse during two months of a single RSV epidemic, reported an average duration of 3.4 –7.4 days (range 1 – 36 days) in 39 otherwise healthy children (Hall *et al.* 1976). Young age, a risk factor for severe RSV disease (DeVincenzo 2005), was correlated with longer shedding duration i.e. children aged less than 2 years had a longer mean shedding duration of 9 days compared to 1.6 days for those 16 years and older (Hall *et al.* 1976). A recent community study involving 193 RSV infected children aged less than 15 years from 151 rural Kenyan families, recorded a mean duration of 4.5 (range 1 – 14) days (Okiro *et al.* 2010). The duration of shedding was not significantly associated with gender, infection severity and age but was reduced in children with a prior history of infection: 4.0 days in children with prior history of infection compared to 5.1 days in those never infected (Okiro *et al.* 2010). In the Kenyan study, follow up to determine the shedding duration started when an individual was identified as RSV positive by immunofluorescent antibody test (IFAT) in nasal washing collected upon signs of ARI. The study design again suffered from bias through possible shedding prior to illness and from excluding asymptomatic individuals. Studies on viral shedding in adults are few. Observations from a study in the USA, reported a mean duration of RSV shedding among 118 infected young adults of 3.9 (range 1 – 7) days. Shedding was detected for less than a week in 93% of the infected adults (Hall *et al.* 2001).

Prolonged shedding of RSV does occur. In the Rochester family study 5 – 10% of individuals shed RSV for more than two weeks and periods of shedding of one month or more have been observed (Hall *et al.* 1976; Waris *et al.* 1992). In South Africa, RSV cases were observed during the inter-epidemic period and this was ascribed to the prolonged shedding in the HIV infected individuals (Madhi *et al.* 2000). Unbiased estimation of the duration of RSV shedding by age, disease severity, infecting group, sex, history of prior exposure, is of importance to the understanding of RSV transmission and persistence in the community, and is focus for the current study.

2.7 Effect of HIV infection on RSV epidemiology

There are limited studies assessing the effect of HIV status on RSV epidemiology. Studies in children with immunodeficiency from other causes in the late 1970s reported a variation in the pattern of shedding (Fishaut *et al.* 1980; Hall *et al.* 1986). In one such study, immunocompromised children shed RSV for >2 times longer and for significantly greater mean peak titers than their normal age-matched controls (Hall *et al.* 1986). The mechanism of the increased duration and amount of RSV shed is not clearly understood. However, there is evidence suggesting that humoral and moreso cell-mediated immunity, which is depressed in immunocompromised children, is key in viral clearance (Fishaut *et al.* 1980). Individuals with HIV infection have poor cell mediated response against viral infections and likely to have a longer duration of RSV shedding compared to the HIV negative counterparts. The impact of HIV infection on the duration and amount of viral load shed has never been evaluated. However, there is indirect evidence from South Africa showing this could differ between HIV infected and non-infected children with a potential of altering the transmission dynamics of RSV infection especially in children. In South Africa, it was reported that (i) HIV infection was associated with 2.5 times greater incidence of RSV admissions (Madhi *et al.* 2001), (iii) the risk for developing RSV-LRTI persists beyond first 6 months of life among HIV infected children (Madhi *et al.* 2001) and (iii) RSV infection occur throughout the year in HIV infected children (Madhi *et al.* 2000). In the current study, the HIV status was not determined among the study infants due to ethical and logistical reasons. The HIV prevalence in coastal Kenya has been estimated to be at 4.2% according the recent Kenya Demographic Health Survey (Kenya National Bureau of Statistics 2010).

Table 2.3: Studies estimating the duration of RSV shedding

Country, Year	Sampling regime	Population Size, age	Specimen, tests	Duration (days)	Comments	Reference
i) Hospital studies						
USA, 1974 – 75	Sampling started after admission and every 1 – 3 days while in hospital and continued at home after discharge until the samples tested negative	23, <2years	Nasal wash, culture	6.7 days; Range, 1 – 21 days	No effect of age and gender. Infants with LRTI vs. URTI (mean 8.4 vs. 1.4 days, p-value<0.01)	(Hall *et al.* 1976)
USA, 1989 – 90	Sampling only during hospital stay. The sampling frequency not defined	41, <4years	NPA, culture and IFAT	40-60% ceased to shed virus 8–10 days after onset of illness	Other factors not assessed	(Waris *et al.* 1992)
ii) Community studies						
USA, 1974 – 75	Samples collected every 3 to 4 days at home regardless of	39, All age	Nasal and throat swabs,	3.4 – 7.4 days Range; 1 – 36	Duration longer for the younger children but most	(Hall *et al.* 1976)

	symptoms	groups	Culture	days	children were positive on first sample collected	
USA, 1975 – 79	Sampling on weekly or biweekly depending on season and year regardless of symptoms. Additional samples collected when ill	44 <4 years	Nasal wash and throat swab, culture	73% shedding during first week and 5.9% in second week	Delay in inoculation of specimen into culture cells. Long shedders (3 children at 18, 24 and 33 days)	(Frank *et al.* 1981)
USA, 1975 – 95	Samples were collected every 2-3 days regardless of symptoms during RSV epidemics	118, 18–60 years	Nasal wash, culture	3.9 days (range, 1–17 days)	-	(Hall *et al.* 2001)
Kenya, 2002 – 05	RSV positives were recruited from family study and sampled at pre-specified days thereafter, and ended upon a single negative nasal wash sample	193 <15 years	Nasal wash, IFAT	4.5 days Range; 1 – 14	No effect of age, gender and disease severity; duration reduced by prior history of infection	(Okiro *et al.* 2010)

Key: NPA, Nasopharyngeal aspirate; IFAT, Immunofluorescence antibody test; URTI, upper respiratory tract illness; LRTI, Lower respiratory tract illness

2.8 RSV treatment

Available therapies for the treatment of RSV infections are limited to ribavirin, intravenous immunoglobulin, and palivizumab. Oxygen supplementation, administration of intravenous fluids and occasionally mechanical ventilation are offered to the severely ill as supportive therapies (Simoes 1999). Though approved by the US Food and Drug Association, Ribavirin use has remained controversial with the American Academy of Pediatrics recommending against its routine use (American Academy of Pediatrics 2012). Randomized controlled clinical trials comparing ribavirin with placebo in children with RSV-LRTI have yielded mixed results. Some studies have demonstrated improvement in illness severity, shortened duration of mechanical ventilation, oxygen supplementation and hospitalisation, and decreased viral shedding (Hall *et al.* 1983; Taber *et al.* 1983; Hall *et al.* 1985; Rodriguez and Parrott 1987; Smith *et al.* 1991), unlike others (Meert *et al.* 1994; Moler *et al.* 1996; Law *et al.* 1997; Guerguerian *et al.* 1999). A Cochrane review of randomized trials comparing aerosolised ribavirin with placebo in infants and children with RSV-associated LRTI found that trials of ribavirin were small and lacked adequate power to offer reliable estimates of the effects (Ventre and Randolph 2007).

Passive immunization with RSV hyperimmune immunoglobulin or Palivizumab, a neutralizing humanized mouse monoclonal antibody directed against the F protein is the alternative in high-risk infants such as those premature, with bronchopulmunary dysplasia or with congenital heart disease (Baker and Ryan 1999; DeVincenzo *et al.* 2003). The prophylactic treatment is administered in the months preceding the start of the RSV season (Hemming *et al.* 1987; Simoes *et al.* 1998; Singleton *et al.* 2006). Despite the observed merits of Palivizumab, high costs have impeded its widespread use. There are no studies looking at the effect of the two approved treatment options on the transmission dynamics of

RSV in the community settings. Ribavirin and Palivizumab are not available for use in low-income settings.

2.9 RSV vaccines

The profile of severe RSV disease suggests a vaccine is primarily required in the paediatric population and high-risk adults such as the elderly and those with chronic pulmonary or cardiac disease. An effective vaccine has, however, remained elusive despite over 50 years of relentless research (Collins and Murphy 2007). The risk of severe RSV disease is highest in infants under 6 months of age and consequently vaccine development in the past has principally been targeted to protect infants in the first few weeks of life. This approach, however, has had several major obstacles: First, the young infants have a relatively immature immune system compared to older children and adults hence are poor responders to vaccine immunogens. Second, the presence of maternal RSV-specific antibodies in newborn seems to interfere with immune responses elicited by the administered vaccines (Murphy *et al.* 1988; Crowe 2001). It is fascinating that RSV infections occur in young infants despite the presence of maternally derived passive antibodies (Ochola *et al.* 2009). Third, infants are extremely susceptible to reaction against even attenuated viruses, used in live-attenuated vaccines (LAV), thus complicating vaccine production regarding balancing between attenuation and immunogenicity (Wright *et al.* 2000; Karron *et al.* 2005). A paradoxical observation of repeat infections by same or similar strains throughout life raises an additional challenge in vaccine development. It appears natural infection confers incomplete immunity to subsequent infections and antigenic variation appears to play only a minor role (Agoti *et al.* 2012). Following infection, protection against re-infection has been estimated to be partial (around 70%) lasting for up to 6 months (Ohuma *et al.* 2012). This protection seems to be dependent on the level of the neutralising antibodies which appear to decay after infection to pre-

infection levels after 3 months {Sande, 2013 #868}. Developing vaccines to offer a superior protection than a natural infection continues to be a hindrance. Heightened safety concerns following the failure of the first RSV vaccine, a formalin-inactivated RSV (FI-RSV), is a further obstacle to paediatric vaccine development. The vaccine failed to protect young seronegative infants against RSV disease. Vaccine recipients experienced enhanced disease after the wild-type RSV infection, resulting in two deaths. The mechanism of the enhanced disease is still not fully elucidated. Lastly, existence of multiple RSV antigenic variants also hampers vaccine development efforts (Durbin and Karron 2003; Nokes and Cane 2008; Medley and Nokes 2009).

During the past decades, several candidate vaccines have undergone clinical trials in humans, reviewed in Table 2.4. These include LAV for intranasal use such as the cold-passaged, temperature sensitive (cpts) RSV vaccine (Wright *et al.* 2007), and live RSV/PIV3 chimeric virus vaccine candidate (Gomez *et al.* 2009; Bernstein *et al.* 2012); subunit RSV vaccines for intramuscular administration such as Purified F protein (Munoz *et al.* 2003; Piedra *et al.* 2003), a purified fragment of the RSV G protein fused to albumin-binding domain of streptococcal protein G (BBG2Na) (Power *et al.* 2001), subunit RSV-A vaccine containing purified F, G, M proteins (Falsey *et al.* 2008; Langley *et al.* 2009) and RSV nanoparticle vaccine candidate (Glenn *et al.* 2013). Other RSV vaccines in early stages of development are epitope based vaccines (Anderson *et al.* 2010) and live vaccine virus attenuated by deletion of non-essential genes or the G-protein (Teng *et al.* 2000; Jin *et al.* 2003; Widjojoatmodjo *et al.* 2010).

Of the vaccines in development, one LAV is promising. This recombinant LAV is well tolerated and immunogenic in young seronegative children aged over 6 months old (Karron *et*

al. 2005). However, in addition to targeting the young infants, alternative strategies for delivery of the vaccine require consideration (Anderson *et al.* 2013). First, by targeting older children (6 – 24 months) who are still at risk of severe disease and probably play a role in spread of the virus in households. Second, to target pregnant women by a maternal vaccine (Munoz *et al.* 2003) in order to protect newborns by passive placental transfer of antibodies and possibly block mother-infant transmissions. Lastly, target the elderly (Falsey *et al.* 2008) and high-risk adults with aim of reducing the disease severity. Deciding on the best strategy or strategies will depend on our further understanding of transmission patterns of the virus in the community. In particular, identifying the groups (e.g. older siblings, mothers or other family members) that transmit RSV to infants for possible targeted immunization is essential. The current study aims to fill this gap and quantify the relative contribution of these groups in RSV spread within household. It is plausible that blocking transmission in the household or reducing the amount of virus circulating in the community can indirectly protect the young infant (<6 months of age). However, to assess the impact of this strategy it is necessary to quantify who acquires infection from whom.

Table 2.4: Status of RSV vaccine candidates

Vaccine type	Platform	Status	References
Live attenuated vaccines	Cold-passage/temperature sensitive	Phase I/IIa	(Karron *et al.* 2005)
	Genetically engineered by reverse genetics	Phase I/IIa	(Wright *et al.* 2007)
Subunit RSV vaccines	Purified F protein (PFP-1, 2, 3)	Phase III-discountinued	(Piedra *et al.* 2003)
	BBG2Na	Phase III-discountinued	(Power *et al.* 2003)
	F/G/M mixed formulation	Phase II	(Piedra *et al.* 2003)
	F/G chimeric protein	Phase I/IIa	(Murphy *et al.* 1989)
Vectored with F or G protein	PIV3	Phase I/IIa	(Schmidt *et al.* 2001)
	Vaccinia virus	Pre-clinical	(Connors *et al.* 1991)
	Adenovirus	Pre-clinical	(Hsu *et al.* 1992)

Key: PIV 3, para influenza type 3; F, Fusion protein; G, attachment protein; M, matrix protein

2.10 Who contacts whom and who infects whom

RSV is transmitted through large nasal droplets and fomites, and has a short survival time in the environment. This suggests that close contact is important for effective transmission (Hall *et al.* 1978). Social interaction patterns (contacts and mixing patterns among individuals) thus have a direct impact on transmission dynamics of RSV. This underlies the growing interest in describing and quantifying contacts that can lead to infection spread (Edmunds *et al.* 1997; Edmunds *et al.* 2006; Mossong *et al.* 2008; Read *et al.* 2008; Vynnycky *et al.* 2008; Zagheni *et al.* 2008). The use of contact data in transmission dynamics modelling requires the inference of the transmission probability following contact – which is a serious obstacle due to its considerable uncertainty. Longitudinal studies of infection and contact can estimate this transmission probability more directly (Melegaro *et al.* 2004).

It is plausible social contexts with high contact rates such as households and schools would provide environments of disproportionate importance for infection transmission. Family studies indicate the importance of household size and school age siblings as risk factors of RSV infection in infants (Hall *et al.* 1976; Okiro 2007). Various studies indicate RSV infection is frequently introduced to the home by siblings, resulting in high secondary attack rates within households (Berglund 1967; Hall *et al.* 1976; Okiro 2007) – although a recent study implicated the young infants (Crowcroft *et al.* 2008). Direct evidence of mother-to-child infection has been reported in one study from Guinea Bissau (Stensballe *et al.* 2004). It remains, however, that patterns of transmission, so important to determine the influence of infection (or control) in one age group on other groups, are poorly defined and will provide a focus for the present study.

2.11 Household studies

Households have been a focus of epidemiological studies on spread of infectious diseases in the past (Longini *et al.* 1982; Viboud *et al.* 2004) and are important in transmission of respiratory viruses due to factors alluded to in the previous sections. Improved understanding of household transmission may inform on the merits of targeted control strategies. For example, investigation of who introduces infection into the household, and who infects the infant, could identify key groups for vaccination such as elder and school age siblings. Data is scarce on intra-family spread of RSV and other viral respiratory infections especially in developing countries. Of the few studies undertaken, most were conducted in 1960s - 70s in developed countries as shown in Table 2.5. The classic family study on RSV spread recruited 36 US families for intensive surveillance over two months in one RSV season, with repeated nasal sampling every 3-4 days regardless of illness (Hall *et al.* 1976). The authors reported an appreciable secondary attack rate within the family of 27%, rising to 45% in infants. The risk of infection was high across all ages in the family: in young children (1- 5 years), older children (5 – 16 years) and adults (≥17 years) the attack rate was 27%, 19% and 17%, respectively. The attack rate was 29% in infants. The proportion of infected older children (43%) and adults (38%) in RSV–exposed families was comparable to young children (46%). These data together with corresponding secondary attack rates of 21% and 33% suggested older children and adults play a role in the RSV spread within families. However, this study had several limitations. First it was initiated after the start of the annual RSV epidemic and hence some families might have experienced the infection before the study started. Second, it covered a short period of the RSV season: it is possible they missed a number of infections occurring before or after the study period. The nasal wash or throat samples were screened using viral culture which is reported to be less sensitive compared to molecular methods

(Falsey *et al.* 2003). Lastly, the study was carried out in the US with demographic features differing from those characteristics of low-income countries such as Kenya. For instance the mean household occupancy was lower compared to rural Kenya (3.5 members versus 8 members in Kilifi, Kenya) (Kenya National Bureau of Statistics 2010). Therefore, it can be asserted that there is need for a longer study covering a full RSV season in a developing country setting with frequent regular nasopharyngeal swab collections regardless of symptoms from all household members and tested using sensitive molecular techniques. The success of such a study will be intimately linked with the acceptability of the specimen collection method in the study population.

Table 2.5: Review of studies assessing RSV infection in households

Location	Study population	Study design	% Infected	Comments
Finland (Berglund 1967)	Family members of child with RSV infection.	Study conducted over one RSV season	Siblings, 23%; adults, 16%	Mothers had higher infection rates than fathers. Children likely introducers
Tecumseh, US (Monto and Lim 1971)	341 families	Families followed for one year noting acute illnesses. Blood samples collected at month 0, 6 and 12 of the follow up	1-4 years, 16%; 5-9 y, 20%; 10-14 y, 17%; 15-19, 10%; 20-29y, 6%; 30-39y, 4%; 40 – 49y, 3%	-
Tecumseh, US (Monto *et al.* 1974)	Six years of data with each family followed up for one year	Families followed for one year noting acute illnesses. Blood samples collected at month 0, 6 and 12 of the follow up	1-4 years, 17%; 5-9 y, 20%; 10-14 y, 17%; 15-19, 10%; 20-29y, 8%; 30-39y, 4%; 40 – 49y, 7%; 50+ y, 7%	-

Rochester, US (Hall *et al.* 1976)	36 families, 188 members	Recruited families with 2 or more, one <1 year. Nose and throat specimens collected every 3-4 days regardless of symptoms for two months.	<1 year, 29%; 1 - <2 years, 29%; 2- <5 years, 26%; 5 – <17years, 19%; 17 – 45 years, 17%	Children the most likely introducer of RSV into the household
Houston, Texas, US (Glezen *et al.* 1986)	Families with 92 children	Birth cohort. Families visited weekly during the respiratory disease season for illness assessment. NW collected at each contact. Serial serum samples collected from the index patient – 4 samples in the first years and 2 spanning each subsequent respiratory disease season	1-2 years, 76%; 2-3years, 45%; >3 - 4years, 33%; >4 – 5 years, 50%.	-
London, UK (Crowcroft *et al.* 1999)	75 Infants admitted in paediatric intensive care units	Study covered one RSV season. Family members were recruited soon after RSV detection in the infants and per-nasal	Children, 54%; Adults, 34%	Families had a median total of 3 members each. Infant and the rest of the family members

	with RSV infections from 74 families	specimen collected for RT-PCR testing		appear to have acquired their infection from someone outside the household
Okiro EA (Okiro 2007)	81 households with 447 members (84 being the birth cohort children) were recruited	Nested in a birth cohort. Recruited family members of the cohort children and followed for two RSV seasons collecting NW specimen when signs of acute illness were observed. Household visited weekly during RSV season and once a month in non-RSV seasons	First epidemic: Birth cohort children, 60%; siblings, 17%; and adults, 0%	Second epidemic: birth cohort, 27%; siblings, 13%; adults, 6%.

Key: RT-PCR, reverse transcriptase polymerase chain reaction; NW, nasal wash

2.12 RSV diagnosis

2.12.1 Specimen type

Nasopharyngeal aspirate and nasal wash (NW) has been the preferred specimen for RSV diagnosis but deep nasopharyngeal specimens collected using flocked swabs have been shown to be a good alternative (Chan *et al.* 2008; Munywoki *et al.* 2011). Nasopharyngeal flocked swabs (NPS) are easier to collect, standardize across age groups and acceptable in a wider range of settings than aspiration or nasal washing. Prior to the present study, we conducted a study to assess the diagnostic performance of NPS relative to NW in outpatient settings (Pingilikani Health Centre in Kilifi District, Kenya) recruiting children presenting with clinical features of URTI such as runny/blocked nose and/or cough, or difficulty in breathing. Using real time M-PCR (the assay used in the current study, see Chapter 3 for details), we found NPS was not inferior to NW in detecting RSV and other respiratory viruses (Table 2.6) (Munywoki *et al.* 2011). However, the mean cycle threshold (Ct) values in detection of RSV were significantly lower (an indication of higher viral yield) compared to NW (P values <0.001. Use of M-PCR increased the prevalence of RSV detection from 14% (by IFAT) to 22%, an indicator of additional benefit of using molecular techniques. In terms of acceptance, most caretakers (60%) and participants (72%) preferred NPS method compared to the NW (Munywoki *et al.* 2011). However, for longer-term studies requiring frequent sample collection, a less invasive specimen such as a pernasal swab or oral fluid (OF) will be required.

Table 2.6: Relative sensitivity of NW and NPS in detection of respiratory viruses using real-time multiplex-PCR among 299 Kenya children visiting outpatient settings with ARI

	Relative sensitivity (95% CI[1])		
Viruses	NW	NPS	**P value[2]**
Rhinovirus	73.4 (62.2 – 82.7)	89.9 (81.0 – 95.5)	0.024
RSV	87.8 (78.2 – 94.3)	94.6 (86.7 – 98.5)	0.146
RSV A	88.2 (72.5 – 96.7)	88.2 (72.5 – 96.7)	1
RSV B	87.5 (73.2 – 95.8)	100 (91.1 – 100)	0.063
PIV	86.5 (71.2 – 95.5)	81.1 (64.8 – 92.0)	1
Adenovirus	45.5 (24.4 – 67.8)	86.4 (65.1 – 97.1)	0.035
Corona virus	86.4 (65.1 – 97.1)	86.4 (65.1 – 97.1)	1
hMPV	86.7 (59.5 – 98.3)	93.3 (68.1 – 99.8)	1
Any virus detection[4]	***79.2 (73.6 – 84.1)***	***89.6 (85.1 – 93.1)***	***0.0043***

Key: 1, one-sided 97.5% CI reported if sensitivity was 100%; 2, Exact McNemar's significance probability values comparing sensitivities for NW and NPS. NPS, nasopharyngeal flocked swab; NW, nasal wash; RSV, respiratory syncytial virus; PIV, parainfluenza virus 1, 2, 3 & 4; hMPV, human metapneumovirus; CI: confidence interval. Table adapted from Munywoki et al (Munywoki et al. 2011))

2.12.2 Oral fluid as an alternative specimen for RSV detection

RSV specific-antibody measurements in sequential sera have previously been shown to be a useful supplement to virus detection methods in identifying infection (Hall *et al.* 1976; Henderson *et al.* 1979; Glezen *et al.* 1986). Okiro *et al* (Okiro *et al.* 2008) evaluated the use of oral-fluid (OF) samples in replacement of serum, finding that RSV specific IgG antibody profiles in OF closely matched those in sequential paired serum samples. However, RSV antibody levels change rapidly and weekly sampling would be ideal to support temporal

resolution of infections (Okiro *et al.* 2008). There is preliminary evidence to indicate reasonable potential for OF samples to detect viral infections using RT-PCR methods, though with lower sensitivity (~70%) relative to nasopharyngeal specimens (von Linstow *et al.* 2006; Robinson *et al.* 2008). The OF method is subject of further investigation in this study particularly as possible alternative in detection of RSV using either ELISA or M-PCR.

2.12.3 Real time multiplex PCR

Molecular methods are becoming more widely used for the detection of respiratory pathogens, in part because of their superior sensitivity, rapid turnaround time, and ability to identify pathogens that are difficult to culture. Over the years, PCR has been the dominant amplification method. However, recently, modifications of this technology have emerged, some of which allow for the rapid detection of multiple pathogens in a single test, like the proposed multiplex PCR. The multiplex PCR is still challenging because amplification conditions for multiple targets are often incompatible and the high concentration of primers typically yields elevated background readings and reduced amplification efficiency. It is therefore important to ensure that a multiplex real-time PCR would not only be cost effective, but also sensitive enough to detect pathogens in multiple infections. Using the samples from the Pingilikani study, we validated the assay used in this study (details provided on the methods Chapter).

2.13 Other respiratory viruses

The advent of molecular diagnostics for a broad range of respiratory viruses has enhanced the study of virus epidemiology due to increased sensitivity and range of pathogens detectable over traditional methods. In the family setting, screening for viruses with modes and patterns of transmission similar to RSV would increase the capacity of the proposed study to identify

household spread and hence improve the definition and quantification of WAIFW for a given sample size. Such data in combination with data on rates of contacts (collected in parallel contact studies – details not contained in this report) is key to validating the use of contact pattern data in predicting the possible spread of emerging viruses (Mossong *et al.* 2008).

In Kilifi District Hospital setting, we have investigated the occurrence of 13 respiratory viruses in nasal specimens collected from 760 severe and very severe pneumonia paediatric admissions throughout 2007 (Berkley *et al.* 2010). RSV was detected in 34% of cases (and 5% of 56 well controls) using real-time M-PCR, compared with 21% of cases using immunofluorescence. At least one virus was present in 56% of cases, with peak occurrence in the months of January-March and November-December, coincident with peak pneumonia admissions. Although longer-term surveillance is required to fully define seasonal trends, we found co-occurrence temporally with RSV of hMPV as previously identified by van den Hoogen *et al* (van den Hoogen *et al.* 2003), PIV3 (in previous studies not always co-circulating) and some coronaviruses (though all at much lower frequency). As an important respiratory pathogen, rhinovirus has elsewhere been reported to be prevalent throughout much of the year and not uncommonly co-circulating with RSV (Phillips *et al.* 1990; Monto 2002; Souza *et al.* 2003; Kusel *et al.* 2006).

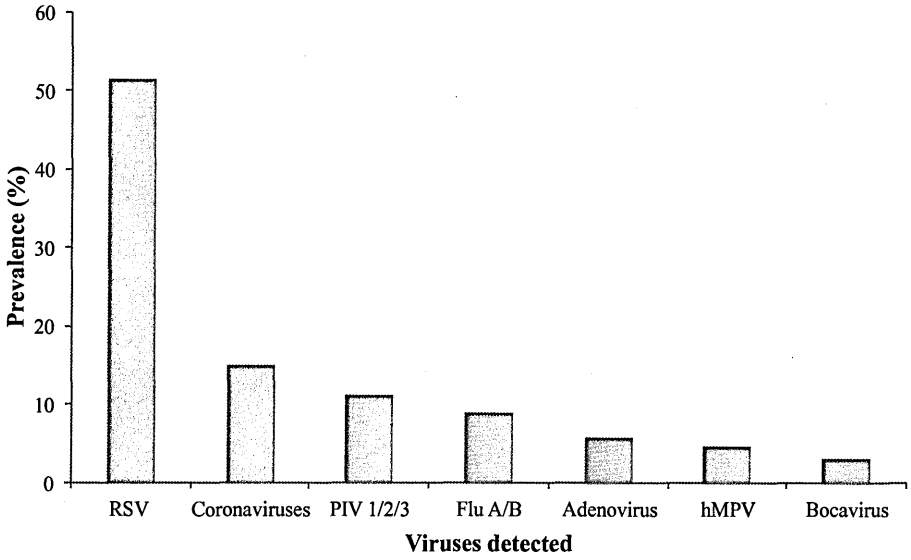


Figure 2.5: Distribution of the detected viruses in 426 children admitted with viral associated severe or very severe pneumonia to Kilifi District Hospital in 2007. RSV, respiratory syncytial virus; PIV, parainfluenza; Flu, influenza; HMPV, human metapneumovirus (adapted from Berkley et al (Berkley et al. 2010).

CHAPTER THREE

3 Materials and Methods

3.1 Introduction

Epidemiological studies in the past have identified households/families as an important element in the spread of infectious disease in the community (Fox and Hall 1971; Longini *et al.* 1982; Cauchemez *et al.* 2009). Members of the same household are thought to have close contacts facilitating spread. Households are also potential social units for delivery of targeted interventions and with the growing interest on alternative approaches in prevention and control of RSV infections and disease, studies in these social units will be informative (Nokes and Cane 2008; Anderson *et al.* 2013). Improved understanding of RSV epidemiology and transmission dynamics in the community should thus involve detailed studies of RSV infection and spread in the households. In particular, there is need for a study to establish who introduces RSV into the household and infects the vulnerable young infant by delineating the chains of transmission within the household. Such a study would require identification of most (if not all) infection chains within the household. In order to identify all infections, frequent sampling even in absence of symptoms from all the household members and screening the samples with sensitive detection techniques is necessary. Given the intensity of such a study, a careful engagement with the local community at all levels and stages of the study would be mandatory.

Previous studies investigating the spread of RSV in families have had several limitations as highlighted in Chapter 2. We designed our study with a view of mitigating most of these limitations. Two studies were pivotal in designing the current study: a family study in the US and another one in Kenya. The family study in Rochester, US, recruited 36 families for

intensive surveillance over two months during the peak of RSV season, with repeated nasal sampling every 3-4 days regardless of illness (Hall *et al.* 1976). The nasal samples were inoculated into viral cultures for detection of RSV. The authors reported most individuals were shedding the virus at the start of sampling restraining the ability to identify who introduces the infection into the families. This pointed to a need to have a clear track of when RSV begins and ends in the community to enable sampling over the epidemic period. Molecular techniques, which are highly sensitive for virus detections, are now available.

The other study was a family study on RSV transmission in Kilifi, Kenya (Okiro 2007). The study conducted by Okiro *et al* aimed at investigating the risk factors of RSV spread and severity. Its methodology, however, had limitations for evaluating who acquires infection from whom (WAIFW) in the household (Okiro 2007). Home visits were conducted once-a-week with nasal washings (NW) collected only when symptoms of acute respiratory illness (ARI) were reported. Screening was by immunofluorescence antigen test (IFAT), which is less sensitive than molecular methods (Casiano-Colon *et al.* 2003; Munywoki *et al.* 2011). Acceptance of the NW method in older children and especially in adults was poor highlighting the need for search of widely acceptable specimen collection methods.

An alternative nasal specimen collection method increasingly used in viral diagnostics was the nasopharyngeal-flocked swab (NPS) (Chan *et al.* 2008). We conducted a study to evaluate the diagnostic performance and acceptability of the NPS against NW in outpatient clinic before the start of the present study. Children (299) under the age of 13 years attending the health clinic in Kilifi District were recruited and paired NPS and NW specimens collected. We demonstrated the NPS was more acceptable to a wider participant age range without loss of sensitivity, relative to the NW (more details in Chapter 2) (Munywoki *et al.*

2011). The samples from the above study were screened using a real time multiplex PCR (M-PCR) assay system on the ABI7500 platform for multiple respiratory virus detection set up with support from the Health Protection Agency, UK. The targets include RSV A and B, influenza (Flu) A (panspecific) and B, parainfluenza (PIV) 1-4, rhinovirus, human metapneumovirus (hMPV), adenovirus and coronaviruses.

The Okiro's et al study also collected oral fluid (OF) samples to detect specific anti-RSV antibodies to supplement infection data. However, the sampling interval was too infrequent (once every 3 months) to identify infections at the required temporal resolution based on antibody profiles (Okiro *et al.* 2008). Studies from our group suggested that levels of RSV specific IgG (and to a lesser extent IgA) in OF track those of serum, but with very rapid post-infection antibody dynamics (Okiro *et al.* 2008), suggesting that weekly sampling could help in identification of infections. In addition, emerging evidence indicate OF samples could be used to detect viral infections using PCR methods (von Linstow *et al.* 2006). The sensitivity of the OF in detection of RSV using molecular diagnostics was assessed in an effort to identify alternative methods to supplement RSV infection data. This was assessed in the current study.

Given the need to detect chains of transmissions within the household, reliance on the occurrence of symptoms to identify infected persons certainly would result in failure to recognize links in transmission events hence need to collect samples regardless of clinical status. Efforts were required to detect all RSV infection which may frequently be mild, of short duration or sub-clinical in older children and adults (Henderson *et al.* 1979). Given a mean duration of RSV shedding of between 3.5 and 9 days (dependent upon age, severity, mode of collection and method of antigen detection) (Hall *et al.* 1976; von Linstow *et al.*

2006; Okiro 2007), frequent sampling intervals would be a prerequisite to avoid a significant loss in detectable cases.

With the chief aim of defining "Who Acquires Infection From Whom" (WAIFW), we mitigated the above limitations by using the most favourable infection detection methods. This included frequent NPS (twice-a-week) and OF (once-a-week) sampling irrespective of symptoms, and use of highly sensitive assays for identification of viruses. Antibody profiling of oral fluid was to provide additional data on infection status. No previous study has combined all these characteristics. Furthermore, we used PCR based methods to genotype the infecting strain providing an added benefit in offering a degree of finger printing of clusters of infection. The aim of the current Chapter is to present details of the methods underpinning this investigation.

3.2 Chapter outline

Detailed descriptions of the study methods are provided in this Chapter. Initially, a general description of the study site and population, together with details on the study design, implementation and data management, are provided. The process of community engagement and sensitization exercises, identification of eligible households, consenting process, data and specimen collection procedures during home and clinic visits, sample processing and testing methods in the laboratory and set data quality control measures are then presented. While detailed results are presented in Chapter 4, results on household recruitment and retention and on sample collection rates are provided here. In addition, temporal patterns of RSV occurrence from the KDH paediatric wards and results from experiments to assess the diagnostics performance of the M-PCR assay are included. The candidate in close collaboration with Professors James Nokes and Graham F. Medley designed the study and

was funded by Wellcome Trust grant (Grant number 090853). The candidate (herein referred to as the project manager) managed day-to-day study activities, and was incharge of a team of fieldworkers, with oversight from Professor James Nokes.

3.3 Study area

The study was conducted in the rural coastal setting of Kilifi County, in Kenya. The County covers a total surface area of 12,610km^2 with a total population of 1,109,735 and around 200,000 households according to 2009 Kenya Population and Housing Census estimates (Kenya National Bureau of Statistics 2010). Kilifi County borders those of Tana River to the North, Taita Taveta to the West, Mombasa and Kwale to the South, and the Indian Ocean is to the East. The study took place in Kilifi District one of the six administrative Districts in the County (Figure 3.1). Kilifi District normally experiences two rainy seasons; long rains between April and July and short rains between October and December. On average, the annual rainfall is about 1,300mm along the coastal belt. It is generally hot (range of maximum daily temperatures 28^{0}C – 34^{0}C) and humid (relative humidity of 70-80%) throughout the year with hottest months in the period of October to April. The main economic activity is subsistence farming, particularly of maize and cassava. Palm wine harvesting from coconut trees is also common in the rural areas. Firewood and charcoal are the common source of fuel, although in a minority of the households paraffin and gas are also used. The main water source is from open or closed boreholes with some areas having piped water (KHDSS, unpublished data) (Scott *et al.* 2012).

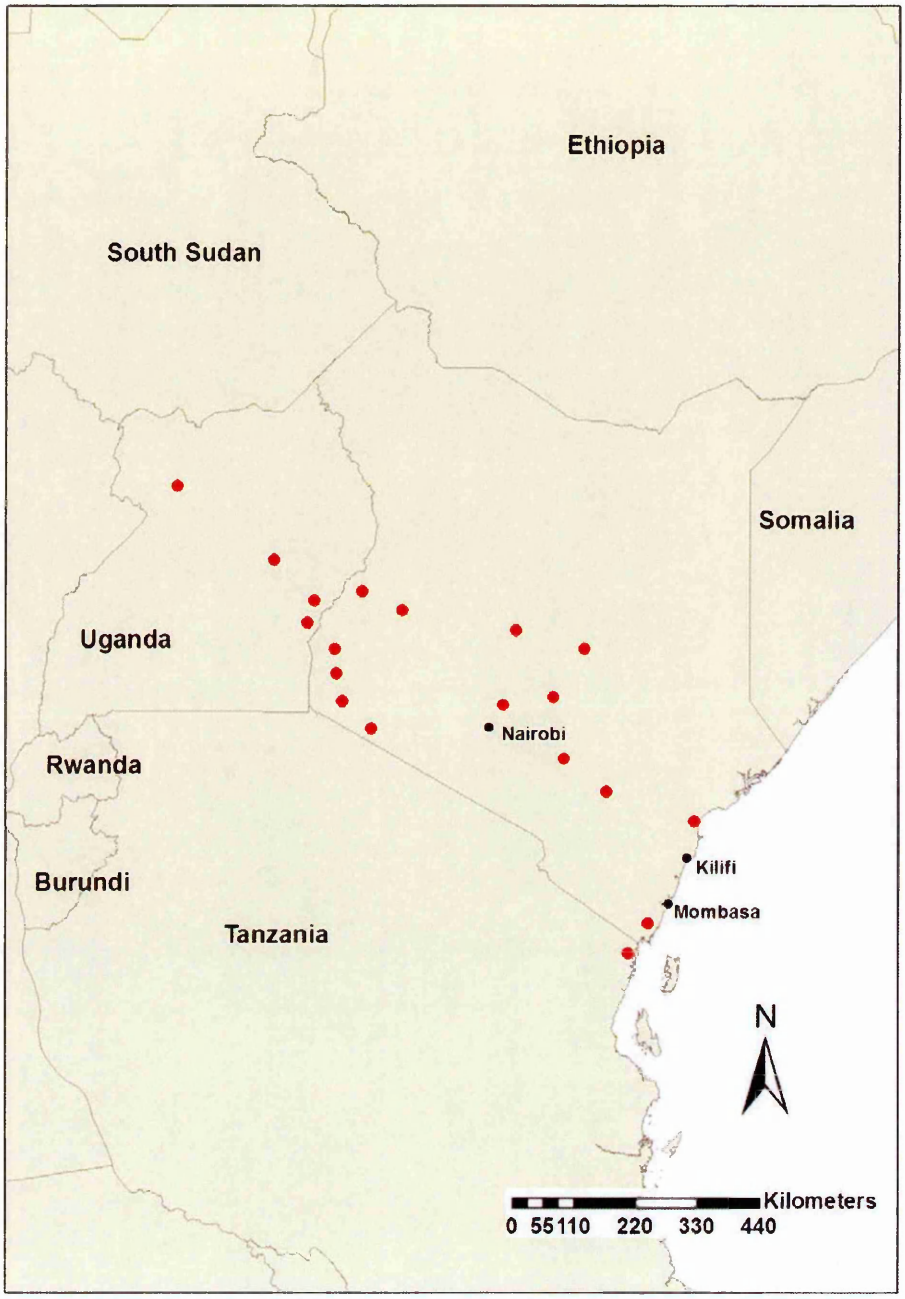


Figure 3.1: Map of Kenya and neighbouring countries. The red dots in the map indicate the location of KEMRI stations. The study was carried out in rural settings along the Coast, north of Kilifi town

3.3.1 The KEMRI-Wellcome Trust Research Programme

The study was carried out at KEMRI-Wellcome Trust Research Programme (KWTRP), which is situated adjacent to Kilifi District Hospital (KDH). The KWTRP is a collaboration between the KEMRI Centre for Geographical Medicine Research - Coast (CGMRC), and the Wellcome Trust, UK. The KWTRP was established in 1989 and runs the children's wards of the District hospital, including a ~60-bed general ward and a 6-bed high dependency unit. In order to facilitate conduct of epidemiological studies, estimate burden of important diseases to the community and evaluate the impact of interventions, the Kilifi Health and Demographic Surveillance System (KHDSS) was set up in 2000 (Scott *et al.* 2012). The KHDSS (light yellow area in Figure 3.2) covers an area of 891 Km^2 and was designed at the outset to capture 80% of the paediatric admissions to KDH (Moisi *et al.* 2010; Moisi *et al.* 2011; Scott *et al.* 2012). Enumeration rounds of all births, deaths and migrations (i.e. into, out of and within) in the KHDSS are conducted approximately every four months. In 2009, approximately 250,000 people were resident in the KHDSS area. Since April 2001, the KHDSS has been integrated with the KDH paediatric wards data (clinical, laboratory and any epidemiological surveillance studies or clinical trials) using a FileMaker Pro Advanced database (FileMaker Inc, US). Furthermore, the KHDSS is linked to an RSV inpatient surveillance study, which has been on-going since 2002 (Nokes *et al.* 2009). The inpatient RSV surveillance study recruits all children aged <60 months admitted to KDH with syndromic severe or very severe pneumonia (Nokes *et al.* 2009). Children are investigated using a nasal specimen (nasal wash (NW) or nasopharyngeal aspirate (NPA)) collected by trained medical assistants soon after admission (Ngama *et al.* 2004) and assayed for RSV antigen by Immunofluorescence antibody tests (IFAT), (DAKO Imagen RSV kit or Chemicon Light Diagnostics DFA kit) (Nokes *et al.* 2009) and additionally by molecular

methods since 2007 (Berkley *et al.* 2010; Hammitt *et al.* 2012). Data from this inpatient study was used in identifying the start and end of RSV season as described later.

KDH is the main inpatient referral facility in the district but there were 23 public health centres and dispensaries offering outpatient services within the District in 2009 (Figure 3.2). The dispensaries provide primary health care; and are staffed by nurses, clinicians, and a public health officer, all with a diploma level training.

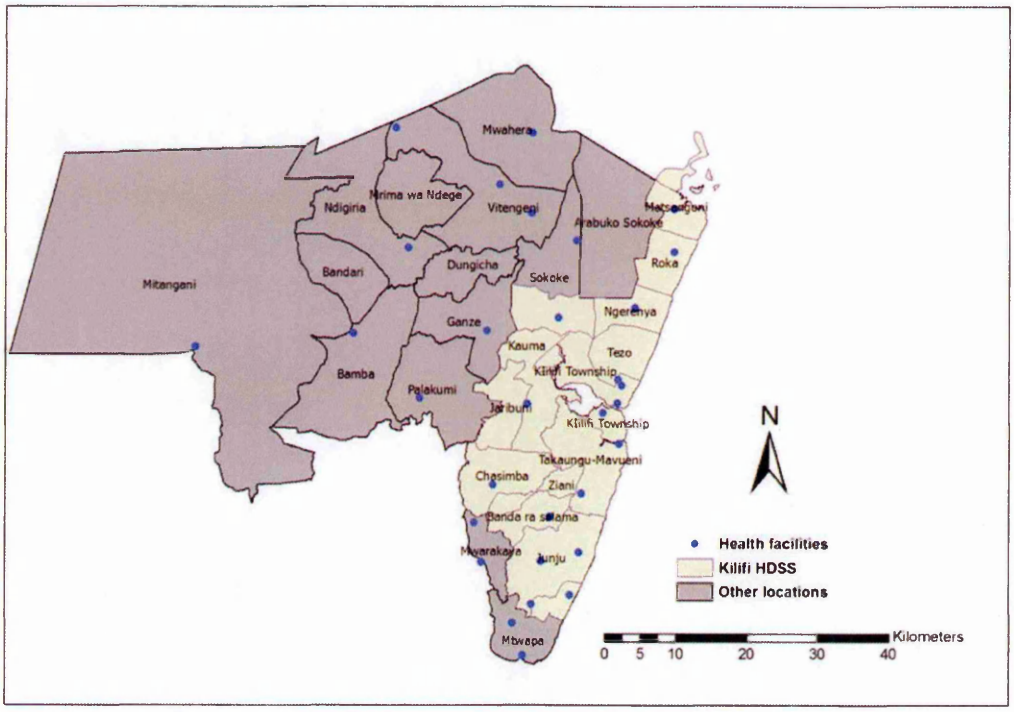


Figure 3.2: Map of the Kilifi District showing administrative locations in KHDSS (light yellow areas) and health facilities (circles filled in red) in the district as at September 2009

3.4 The population

Kilifi is predominantly inhabited by members of the Giriama comunity, one of the nine sub-ethnic groups of the Mijikenda people found along the Kenyan coast. In addition to local Kigiriama, Swahili language is well established in the region. The community is culturally diverse with various religious practices (Islamic, Christian and traditional).

3.5 Methodology

3.5.1 Study design

A household-based prospective cohort study design was used. Households from the KDHSS register were systematically selected from within a single administrative location until 50 households, which fulfilled the eligibility requirements, were identified. Eligible households had an infant (hereafter referred as the study infant) born since the previous RSV epidemic (i.e. born after 1stApril 2009) and one or more siblings <13years old. A household was defined as the group of individuals living in the same compound and eating food from the same kitchen, regardless of their relationships with one another (Table 3.1). The target households had at least one other child potentially increasing the probability of RSV introduction in the household. We aimed to recruit all the members from each eligible household prior to the start of the 2009/2010 RSV season (Figure 3.3). RSV season was formally defined in Table 3.1. RSV seasonality (start and end of RSV epidemic) was identified from our long-term KDH paediatric RSV surveillance (Nokes *et al.* 2009). From the start of the 2009/2010 epidemic, NPS for virology and standard illness assessment were requested every 3-4 days from all individuals, irrespective of symptoms, until the end of the RSV season. Oral fluid (OF) (also referred to as saliva) samples were collected weekly, during a swabbing visit, to investigate OF as an alternative to nasal sampling for respiratory virus detection using M-PCR assay and to assist in identifying infections serologically. A

piloting phase preceded full sampling, in which participants were visited 'once-a-week' for about four weeks. The pilot phase aimed at establishing the household routines and best times for home visits as well as confirming the household members' willingness to participate in the study. The flow of the study processes were streamlined during this phase. Households lost to follow-up during the early phase of the study were replaced by enrolling new households. Newborns into the study households were recruited as individuals who became residents during the course of the study (note that even so there remained only one study infant per household).

i) A summary of the inclusion criteria:

a) Household located within Matsangoni location and registered in KHDSS

b) Household with a child born after 1st April 2009 and at least one sibling <13 years old.

c) Household in which agreement to participate was established for all members who were resident at start of study i.e. envisaged to be living in the location for at least the next three months

ii) Summary of the exclusion criteria:

a) Household with at least one individual refusing to participate at the outset

b) Household in which one or more members withdrew during the study pilot phase

c) Household members who emigrated from the study area in the course of the study

Table 3.1: Definition of terms

Term	Definition
Household	A group of individuals living in the same compound and eating food from same kitchen
Study infant	The youngest child in the household at the time of the recruitment. All the study infants were born after 1st April 2009 and considered RSV naïve.
RSV season	The 'periods delimited by weeks in which one or more RSV cases were identified in our hospital surveillance and within which at least 3 RSV cases were found in any contiguous 3-week period' (Nokes *et al.* 2009)
Away	Status recorded if an individual was out of the household for more than 3 consecutive days e.g. the individual having travelled or in a boarding school
Visit	Includes the instances where the field workers formally met the study participants either at home or at the study clinic verified by filling of the home or clinic visit form. This also includes records of missed appointments for instance, when the participants were away
ARI	Presentation with one or more of the following symptoms: cough, runny nose/blocked nose, or difficulty in breathing
Individual episode or RSV infection episode	Period within which an individual provides specimens which are PCR positive for the same infecting RSV group with no more than 14 days separating any two positive samples. If an individual was diagnosed with both RSV group A and B in the first sample of the individual episode this was coded as a coinfection and counted as one individual episode. Individual episodes are also referred as RSV infection episodes in the text.

Key: ARI, acute respiratory illness; PCR, polymerase chain reaction

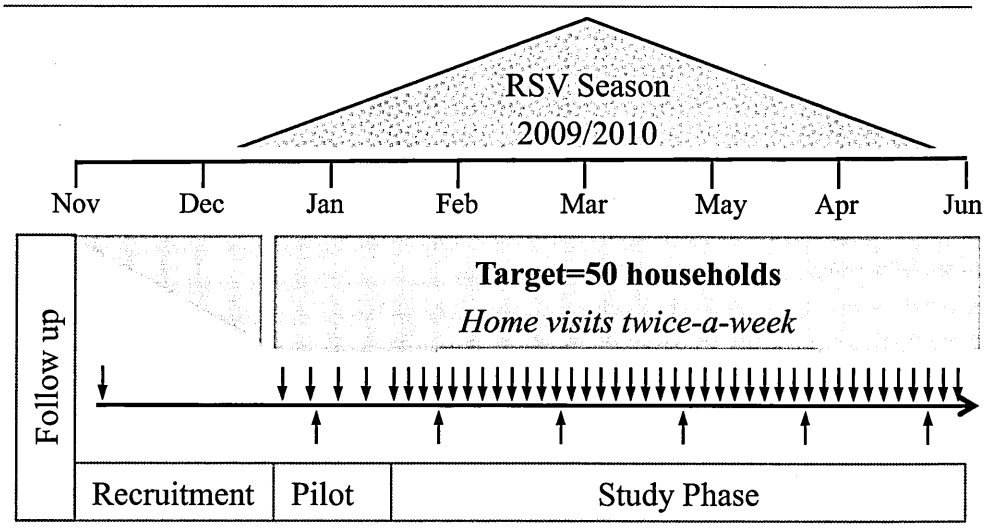


Figure 3.3: Schematic diagram showing the household study design. The dotted triangular represents rise and fall of RSV cases throughout epidemic. Light grey area shows the recruitment and active home follow-up phases; Arrows (black) indicates the regular active home visits while red marks indicate possible passive clinic visits (see section 3.5.1 for details)

3.5.2 Sample size estimation

The unit of sampling was the household. Previous community studies indicated 30-70% of newborns are usually infected during their first epidemic season (Hall *et al.* 1976; Glezen *et al.* 1986; Nokes *et al.* 2004). Our estimate, from the Kilifi birth cohort, (Nokes *et al.* 2004) of around 39% was based on less sensitive methods and was likely to be much higher in the current study – which used molecular virus detection methods. Our primary outcomes were related to RSV infections. We aimed to estimate the proportion of households (i) with at least one occupant infected with RSV at the end of the RSV season, (ii) in which infants experienced a primary RSV infection and (iii) in which infants were infected from within (elder siblings or parents) and from outside the immediate household. Using standard sampling theory (Smith and Morrow 1992; Kirkwood and Sterne 2003) the precision

estimates of the possible proportions for each of the above outcomes at varying sample sizes are shown in the Figure 3.4 below. For instance if the proportion of households with detected RSV infection is 30%, this would give a precision of +/- 13% for a sample of 50 households. We thus set up a prospective cohort study of 50 households (estimating that in each household had an average of eight members), to be followed for one complete RSV epidemic.

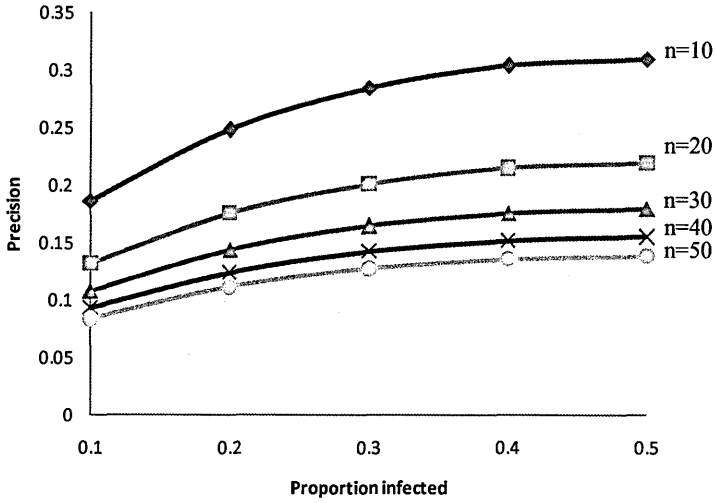


Figure 3.4: Precision estimates of the risk of infection at 5% significance level (two-sided) for specified sample sizes of households (n)

3.5.3 Sampling frequency estimation

The following rationale was used to determine the NPS collection frequency. Assuming individuals shed virus for a mean duration of between 3.5 and 9 days (Okiro *et al.* 2010), with a constant rate of recovery from shedding, and an onset on average half way between any sampling interval, then the proportion of individuals predicted to remain shedding, and thus detectable, would range from 61 – 82% (for 3.5 – 9 days duration) for a 3.5 day sampling interval (Figure 3.5 and further details in Appendix B). The comparable range was

37 – 68% for a 7-day interval in sampling. Given the need to detect infection in mild cases and in older children and adults with likely lower range of shedding duration, sampling twice weekly was preferred. While frequency of sampling greater than twice weekly would undoubtedly yield higher detection rates this was deemed unacceptable for reasons both logistic and to do with participant acceptability. The option of less frequent sampling but inclusion of more households was deemed insensitive in detection of RSV transmission events within the household and suboptimal in addressing the study objectives.

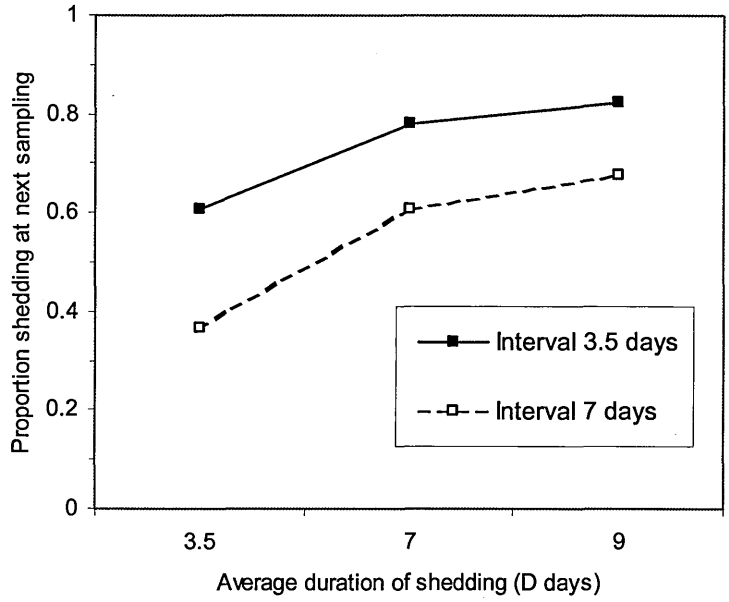


Figure 3.5: Estimated proportion of individuals continuing to shed RSV at the next sampling by average duration of shedding for the two sampling intervals i.e. twice- and once-a-week

3.5.4 Choice of the study site

Due to the intensive nature of the proposed sampling regime and for practical reasons, the study was conducted in one administrative area. The aim was to select a locality fulfilling the following requirements:

a) Relatively easy to access. The roads to be accessible even during the rainy season and with journey time of less than 1 hour from KWTRP in Kilifi
b) Availability of active Community Health Workers (CHW) charged with specific households, to allow easy identification of eligible households and help in community entry
c) Evidence from the ongoing hospital surveillance that RSV recurrently occurs in the location. All the locations within the KHDSS were suitable based on this criterion (Figure 3.6)
d) Availability of a nearby health facility. For participant referral and to set up a study clinic and office for the field team
e) Acceptability of the study by the community leaders

The choice was made after visiting several sites and holding discussions with community leaders including KEMRI-Community Representatives (KCR)[1] (Marsh *et al.* 2008), health facility staff, and local administration in the respective locations. In these meetings, the details of the proposed study were presented and community leaders allowed to ask questions related to acceptability of the study. A leaflet with study details (Appendix C) and another one with list of frequently asked questions (Appendix D) were also shared in these meetings.

---

[1] *The KCR are community leaders voted in by the community members as part of KWTRP community engagement policy in the KHDSS. The KCR meet with members of Community Liaison Group (CLG) three times a year at location level to get their views about KWTRP and its activities, address their concerns and pass any information KWTRP may have to the community.*

Matsangoni location, 20 km from Kilifi town in the northern part of the KHDSS was the only location with 'active' CHW and fitted well with the other requirements. This is a rural location with a population of predominantly subsistence farmers and is under regular health and demographic surveillance. This location has three administrative sub-locations of which two East of the main Kilifi-Malindi road were chosen i.e. Uyombo and Matsangoni (see Figure 3.4).

The location had a total population of 14,998 in 1,835 homesteads[2] in 2009 (KHDSS unpublished data, 2009) distributed as shown in Table 3.2.

---

[2] *Homesteads comprise of a group of individuals living in the same compound. They may consists of one or more households*

Table 3.2: Population distribution in Matsangoni location based on KHDSS data as at July 2009

Sub-locations	No. of homesteads	Population	Individuals per homestead[1]
Matsangoni[2]	544	4513	8.3
Uyombo[2]	624	4984	8.0
Mkongani	667	5501	8.2
Total	***1,835***	***14,998***	**8.2**

Key: 1, average number of individuals per homestead; 2, the study sub-locations

Table 3.3: Homesteads in Matsangoni location based on the Kilifi Heath and Demographic Surveillance System (KHDSS) data as at 20th July 2009

	Number of homesteads			
Homestead composition	Matsangoni	Mkongani	Uyombo	Total
All homesteads	544	624	667	1,835
Homesteads with at least one infant	66	70	71	207
Homesteads with an infant and another child <13 years	65	65	65	195

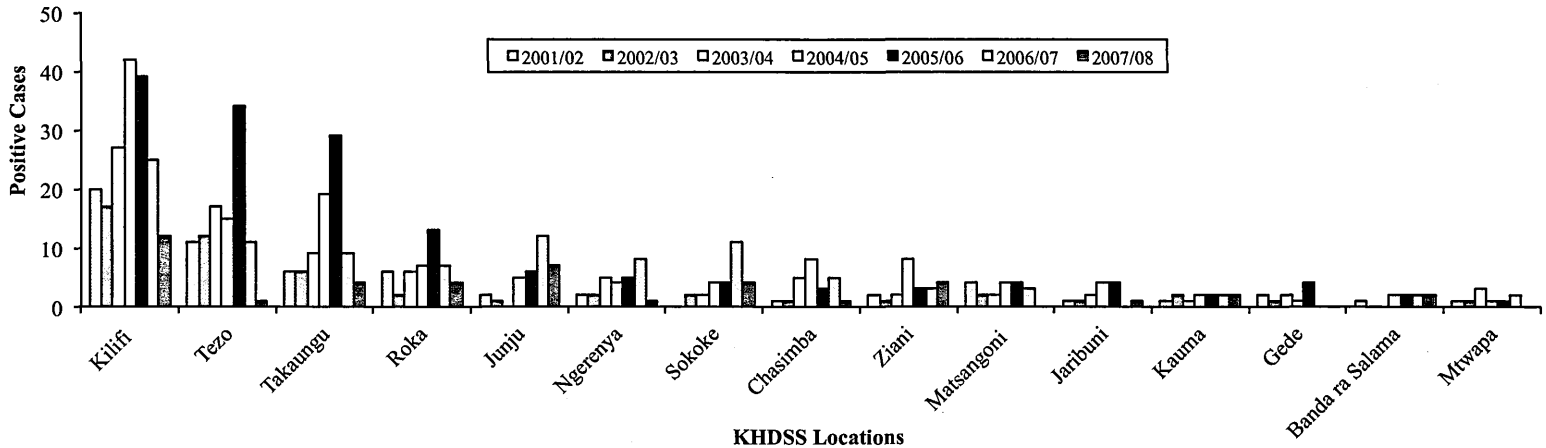


Figure 3.6: Frequency distribution of RSV inpatient cases at Kilifi District Hospital paediatric wards from 2001 to 2008 by administrative location in KHDSS

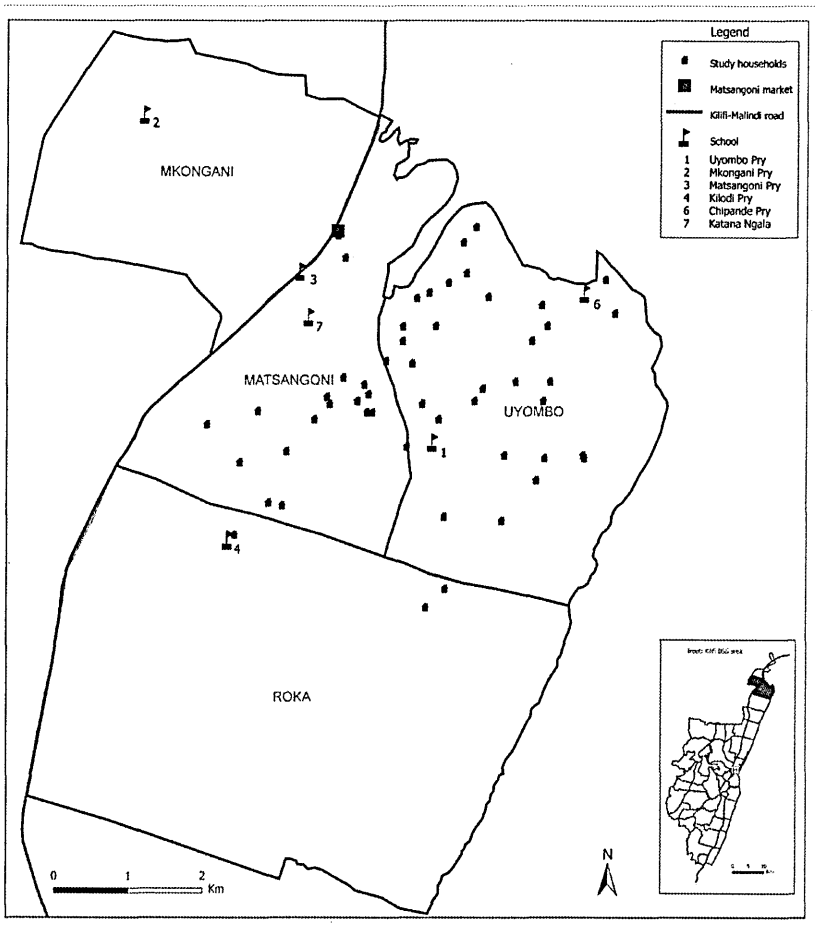


Figure 3.7: Matsangoni map showing distribution of the recruited households (navy blue). Insert at top-right shows the legend and on the bottom-right is the map of the KHDSS area with the grey region showing Matsangoni location

3.6 Study implementation

3.6.1 Recruitment of field staff and training

A careful selection and thorough training of staff was obligatory, given the intensity in sampling and community interactions in the study. Recruitment of the field staff was done in September and October 2009, about two months prior to start of the study to allow proper training. The underlying principle in these appointments was to build a field team with knowledge of the local area and who spoke and understood the local languages and would

easily adapt to the study households' daily routines for ease of sampling. Initially, five field workers, two data clerks, a senior field worker, a study coordinator and clinician were hired. Later in March 2010, two field workers and three community-based assistants were recruited when the full scale of the workload became apparent. Figure 3.8 shows the composition of the study team. The candidate took a lead role in these appointments – creating the job description, advertisement of the positions, preparing the shortlist and interviewing. The field workers and community-based assistants were tasked with conducting the home visits for specimen and data collection and were under direct supervision of the senior field worker. The study clinician assisted by the study coordinator conducted regular (once every month) home visits to carry out health checks on family members and *ad hoc* home visits to discuss any arising issues related to the study participation. The field workers were assigned specific households and the respective field workers were involved in these *ad hoc* visits and on most occasions only aimed at addressing a specific issue. During the early phase of the study, the candidate (the project manager) always accompanied the clinician in these visits.

Prior to start of the study the field team was trained on the following areas, part of a well developed curricula (Appendix E) for new field staff in the programme: -

a) KEMRI and its roles
b) What is medical research & medical ethics, i.e. the protection of participant rights in research
c) Communication skills including problem solving skills
d) Information giving and consenting, involving role-plays simulating home visits and the consenting process
e) Sample collection procedures, nasal swabbing, using nasopharyngeal flocked swabs (NPS) and oral fluid collection

f) Identification of clinical symptoms especially respiratory related, i.e. cough, running nose, counting respiratory rate, heart rate etc

g) Household study details, i.e. Study objectives and implementation process including conduct of home visits for sample collection and referral process to the dispensary

h) Data collection procedures including reviewing of all the data collections forms and piloting

The details of the above training were summarised in a study manual issued to all the field staff for reference (see Appendix F). Refresher training on communication and study procedures was carried out intermittently in the course of the study.

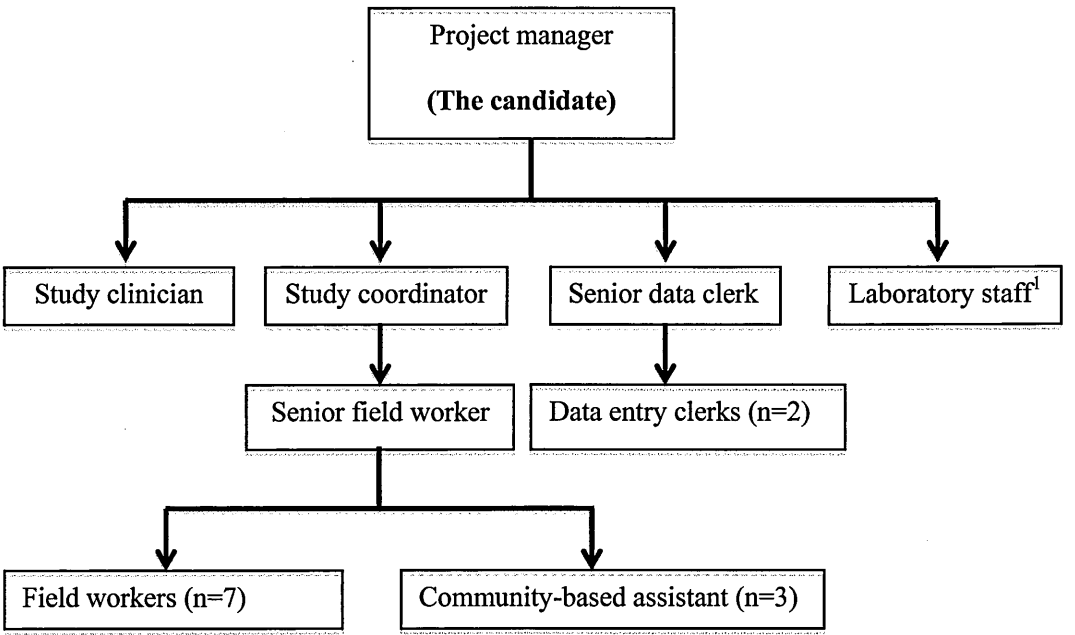


Figure 3.8: An organization chart of the household study team. 1, the laboratory technologists were under direct supervision of a laboratory manager but for the household study sample screening the candidate scheduled their daily activities

3.6.2 Community engagement and sensitization

A detailed community engagement plan was drawn up at the outset of the fieldwork. This was in consultation with a 'communication advice for specific studies' (CAST) group which is spearheaded by the centre's Community Liaison Group (CLG) (Marsh *et al.* 2008). As part of this plan, we held several (i) consultative and sensitisation meetings with the local administration and community leaders or their representatives at various levels from the district to the village (Figure 3.9), (ii) public meetings *i.e. barazas* for sensitisation, and (iii) meetings with the household heads in the area. This helped to establish proper communication channels and foster a favourable environment in the community as well as to obtain community consent. In particular, we could receive concerns easily from the community members through the CHW or other community leaders. The Ministry of Health (MoH) was engaged from the District Medical Office, including the District Health Management Team, to the dispensary level. In particular, the Matsangoni Dispensary Health Committee (DHC) was deeply involved in the sensitisation exercise in view of its pivotal role in the delivery of health care services in the community. The study team was vigilant and proactive in identifying any community fears or problems, which were duly discussed and addressed with the concerned parties. At least once a month, we shared with the local administration on the progress of the study and addressed any concerns raised through their office. Feedback of the study progress and findings were made periodically to the community leaders during their monthly developement meetings at the location. At the end of the field work a comprehensive feedback was made involving local administrative officers, the CHW, various community leaders and health facility staff. The final feedback meeting was held on 14th November 2012 once most of the study findings were identfiied.

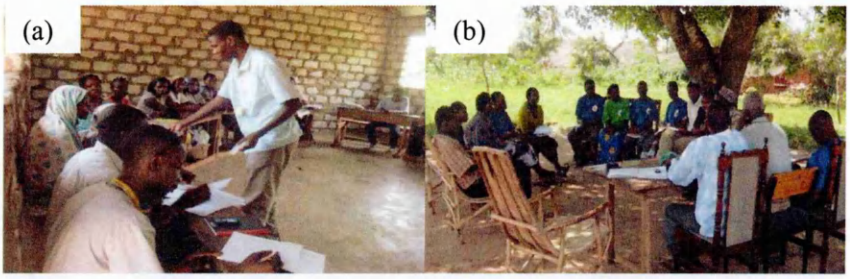


Figure 3.9: Pictures of community sensitization meetings. (a) The project manager (standing) demonstrating OF kit during a meeting with household heads in Uyombo primary school; (b) a meeting with the KEMRI-Community Representatives outside Matsangoni Chief's office

3.6.3 Household selection and recruitment

Eligible households within Uyombo and Matsangoni sub-locations were initially identified using the KHDSS registers. The household lists for each sub-location were subsequently reviewed in consultation with the local CHW and the field workers who were privy to the household details. This resulted in the addition of a few households, particularly those who had newborn not yet enumerated by the KHDSS staff. Field workers visited the potential households accompanied by local CHWs as part of a household entry plan. The aim of this first home visit was to (i) confirm eligibility, (ii) notify the household head about the planned study, if they were not yet aware and (iii) invite the household heads or representatives to a joint sensitization meeting in the nearby school. Upon sensitization, the field workers arranged a second home visit for individualized information giving and for consenting (see Appendix G). In line with local practice, the field workers were again introduced to the household members by the CHWs but the CHWs were not involved in the consenting process. Information was delivered in the local language and comprehension assessed through a set of questions on key messages of the study. Members were given adequate time (within one week) to discuss amongst themselves before deciding whether to participate in the study and in most instances consent was not given by all household members on the same

day. Some required more time to make the voluntary decision to participate. Individual written informed consent was sought from each of the household members. Consent for children under the age of 18 years was sought from their parents or guardians. Verbal assent was also sought from children aged 5 – 17 years with particular focus on the ensuring the teenagers understand and voluntarily agree to participate in the study. Households were excluded if one or more individuals declined to give consent even if some individuals had already consented. Upon consent, the field workers collected baseline demographic and living arrangement data using an Initial Home Visit Form (Appendix H). The contents of the form are described in the data handling section. Household recruitment started in Uyumbo area and progressed into Matsangoni sub-location until the required sample size of 50 households was achieved.

3.6.4 Household follow-ups

Sampling was scheduled to start when the KDH in-patient paediatric surveillance showed two successive weeks with more than two children admitted with RSV positive infections, signalling the start of the RSV epidemic (Nokes *et al.* 2004). For the first three weeks, home visits were made once-a-week but increased to the stipulated frequency of every 3-4 days (nasal swabbing in all the visits and OF in one of them) thereafter. The early phase of sampling allowed the study team to develop a logistical framework/system synchronised with the community activities and culture and replace non-compliant households before the epidemic peaked in the local community. During each visit, a Home Visit form was filled. The form, described in detail later, captured data on samples collected and presence of respiratory symptoms (see Appendix I). Field workers made arrangements to revisit participants who were not available during a scheduled home visit and at least three attempts were made. Participants who were 'away' from the home for more than three consecutive

days were noted. During the field workers' home visits, participants with fast breathing for age and/or fever or any other clinical features of severe illness were referred to the Matsangoni dispensary for medical care, which was provided at no charge at the study clinic. Self or passive referral was also encouraged. At the dispensary, a detailed clinical assessment was made and documented in a Clinic Visit form (Appendix J for details) by the study clinician. Data on a range of household characteristics (e.g. income, education level, housing, anthropometric measures) was also collected towards the end of the study, when a good rapport had already been established with the families, via an interviewer-administered questionnaire (see Appendix K). Home visits (and specimen collection) were terminated when two or more household members repeatedly failed to comply with the sampling regime (i.e. failure to collect samples for a period of more than 4 weeks) or the RSV 2009/2010 epidemic ended (≤2 RSV infections detected from the KDH inpatient surveillance of paediatric pneumonia in two successive weeks).

3.6.5 Study monitoring

A field office and study clinic was set up at the Matsangoni dispensary (see map in Figure 3.2 and 3.7). The field workers had access to the field office even during weekends and outside working hours on weekdays. Every day the field team reported to the office before departing for the home visits. All study supplies including data collection forms and sample collection devices were available from this office. The study coordinator replenished stocks regularly from KWTRP main stores. The field workers kept a log of the daily activities in their individual 'field' books. A main sample collection and reception logbook was kept in the office, which was updated on a daily basis noting when samples were received at the study office and transported to KWTRP laboratories. A separate logbook existed for recording reception of completed data collection forms and when they were forwarded for data entry.

The field team including data clerks and laboratory staff had regular weekly meetings to discuss the study progress. The project manager chaired these meetings. The agenda included a review of the work done and experiences in the preceding week. Any concerns from the study participants or the community were raised and discussed. In instances where matters related to community engagement were to be discussed, field officers from the CLG were invited. This forum was also used for continuous training of the field team on sample collection and data collection. On average, the meetings lasted 2 hours.

The project manager randomly selected days to accompany the fieldworkers during the routine home visits for sample collection and illness assessment. The participatory supervision helped in early identification of any issues that required attention of the project manager such as difficulties in sampling some of the participants and was useful in tracking of the study progress.

3.7 Sample collection and handling

3.7.1 Specimen collection

A description of NPS and OF specimen collection procedures is provided below.

i) Nasopharyngeal flocked swab (NPS)

The method adopted, described in Chan et al (Chan *et al.* 2008), was rigorously tested in our setting for sensitivity and acceptability before use in the current study (Munywoki *et al.* 2011). This is a simple procedure using a commercially available nasopharyngeal flocked swab (Copan Diagnostics Inc, Brescia, Italy). Briefly (see Appendix L for the detailed standard operating procedure (SOP), the distance between the participant's nares and earlobe was measured to estimate the length of insertion. The swab was then gently inserted up the nostril towards the pharynx for the measured distance. The swab was rotated 3 times, to obtain epithelial cells and surface colonising viruses and held in place for 5 seconds to allow

absorbance. The swab was then withdrawn gently and placed in tube with 1 ml of viral transport medium (WHO 2006).

ii) Oral fluid (OF)

Oral fluid was collected using a sponge swab (Oracol, Malvern Medical Developments, Worcester, UK), consisting of a cylinder of expanded polystyrene foam attached to a plastic stick and was used like a toothbrush (Nokes *et al.* 2001; Okiro *et al.* 2008). The swab was brushed along the gums and mouth for 60 seconds and the device was then inserted into a plastic tube and stoppered. The SOP is detailed in Appendix M.

3.7.2 Specimen handling, transportation and storage

Upon collection, samples (NPS and OF) were stored in a cool box (with ice-packs) and delivered to the field office within 4 hours in Matsangoni dispensary for storage in a 4°C gas-powered fridge. Samples were transported to the KWTRP virology laboratory in Kilifi within 24 hours for processing and storage. In the laboratory, 1ml of preservative buffer (20% foetal calf serum and 0.2% sodium azide in phosphate buffered saline (PBS)) was added to the OF sample. The oral fluid was then squeezed out of the foam and clarified by centrifugation and the supernatant stored. Details on the freezer position were appended to the specimen tube before storage of both specimen types in a -70°C freezer for later testing.

3.8 Laboratory procedures

The following describes the procedures for (a) identifying the presence of specific viral nucleic acids in NPS samples (b) serological testing of OF for RSV specific IgG and (c) detecting RSV specific nucleic acid in OF.

3.8.1 RNA extraction methods

Viral RNA was extracted from NPS using MagNA Pure LC RNA Isolation Kit - High Performance (HP) following manufacturers instructions (Roche Diagnostics, Mannheim, Germany). This is a semi-automated method of RNA extraction using the MagNA Pure LC Instrument. The performance of the HP extraction kit was initially compared to MagNA Pure total nucleic acid (TNA) isolation kit and manual extraction by QIAamp Viral RNA Mini Kit (QIAGEN, Valencia, CA, USA). The main aim was to assess for non-inferiority of the HP kit against the Qiagen manual extraction in detection of RSV.

3.8.2 Multiplex real-time polymerase chain reaction (M-PCR) assay

We used a multiplex real-time polymerase chain reaction (M-PCR) assay to detect a range of respiratory viruses. The M-PCR assay employed uses dual-labelled hydrolysis probes to generate the positive signals (Gunson *et al.* 2005). The assay consists of five triplex reactions and one single reaction and detects specific nucleic acid sequences of the following pathogens: RSV group A and separately group B; influenza virus types A, B and C; parainfluenza (PIV) virus types 1 to 4; human coronaviruses (HCoV) strains 229E, OC43 and NL63; adenovirus; human metapneumovirus (hMPV) groups A and B; *Mycoplasma pneumoniae*; and human rhinovirus. Although the hMPV primers for group A and B are present, the fluorescence of the probe specific to each target are close in wavelength and cannot be distinguished. The West of Scotland Specialist Virology Centre originally developed the M-PCR assay and kindly provided the relevant standard operating procedures (SOPs). The assay has been successfully implemented in KWTRP virology laboratory (Appendix N) as a diagnostic and epidemiological tool with support from the Health Protection Agency, Centre for Infections, UK and has been used in a number of studies (Hammitt *et al.* 2011; Munywoki *et al.* 2011; Hammitt *et al.* 2012). The real time M-PCR

assay used undergoes 40 cycles of amplification and a cycle threshold (Ct)[3] value of 35 or less was considered positive.

In order to further examine the performance of the M-PCR assay in virus detection, we (i) compared the sensitivity of a uniplex (single pathogen detection) versus triplex (detecting three pathogens in a single tube) for RSV A/B and adenoviruses, and (ii) assessed the Ct profile of multiplex PCR in IFAT positive and IFAT negative samples. The latter was aimed at assessing the nature (by viral load terms) of any additional sensitivity by our M-PCR assay. In this experiments, we used nasal samples from a previous outpatient study conducted by the candidate (Munywoki *et al.* 2011).

3.8.3 Attachment (G) Gene sequencing

In a select number of samples (details on the selection criterion provided in Chapters 4 and 5) the sequencing of the long ectodomain region of the RSV G gene was carried out as previously described (Agoti *et al.* 2012). This sequencing work was done in close collaboration with Charles Agoti, a PhD student in the Viral Epidemiology and Control (VEC) group interested in molecular diversity of circulating RSV viruses.

3.8.4 Sample screening strategy

A large number of specimens (NPS and OF) was expected, necessitating a well-defined screening strategy to ensure the screening kits were used most efficiently to achieve the

---

[3] *Ct values are inversely proportional to the amount of target nucleic acid in the sample i.e. the lower the Ct value the greater the amount of target nucleic acid in the sample*

desired goal of identifying who infects whom within the household. The project focused on transmission of RSV and all samples were screened for RSV A and B. Only the most prevalent of the remaining respiratory viruses were targeted for screening in all the samples.

i) NPS screening:

In order to identify the common viruses present during the study period, samples from six households were screened for the full range of the respiratory pathogens included in the M-PCR assay. The relevant components of the M-PCR assay were then selected to cover detection of the most prevalent viruses for the remaining sample set.

ii) OF screening with M-PCR assay

To assess the sensitivity of oral fluid relative to NPS in detection of viral antigens, a selected number (60) of samples were tested by the M-PCR assay. The OF samples were selected randomly from a list of samples whose paired NPS (collected on the same day) was PCR positive. Principally this was for RSV but to be extended to cover other viruses if a promising sensitivity (>80%) was reported for RSV detection.

iii) OF screening with RSV-specific antibody assay

RSV specific antibodies in OF samples were quantified by optimised indirect enzyme-linked immunosorbent assay (ELISA) established in our laboratory that uses crude RSV A2 lysate (Okiro *et al.* 2008). Interpretation of changes in specific antibody concentration and the definition of re-infection were made based on the antibody profiles as described elsewhere (Okiro *et al.* 2008). It was envisaged that older individuals especially adults would have poor PCR detection, due to low viral loads and that the ELISA data would supplement the RSV infection data.

3.9 Data collection, handling and analysis

3.9.1 Data collection forms

Field data were collected using four different forms, which have been highlighted previously. These include Initial Home Visit form, regular Home Visit form, Clinic Visit form and Household Risk Survey Questionnaire.

i) The Initial Home Visit Form

This form was developed before the start of the study by the project manager and administered immediately after the household consenting. The form captured baseline demographic characteristics at household and individual level. Data collected included household head name, his/her marital status, and highest educational level and number of families in the household. In addition, participant specific details such as date of birth (mainly from birth certificate or national identity card), occupation, education status, relationship to the study infant, as well as living arrangement (live or sleep in same house as the study infant) were collected. The field workers collected these data.

ii) Regular Home and Clinic Visit Forms

The two forms were developed and piloted during the early phase of the study. Following feedback from the participants, field workers and clinician administering the forms the questions were revised. For the Home Visit Form, which was filled twice a-week for every participant, the priority was to make the data collection exercise brief and targeted. Thus the form was restricted to collecting data on specimen (NPS and OF) collection, quick illness assessment for presence of respiratory symptoms (cough, runny nose/blocked nose, difficulty in breathing etc.) or other complaints and recording of vital signs such as temperature and respiratory rates for the under five year olds. Reasons for not collecting specimen or any other complaints were also captured. For ease of data collection each household had a customised form with the list of members already included and the required data on a tabular

format on one side and allowance for entering additional comments on the back of the form. The field workers were required to record either 'yes or no' response for most of the questions though in some instances a value record was required. The Clinic Visit form was similar to the Home Visit form except that the illness assessment was elaborate. The form was adopted from earlier studies (Okiro 2007; Munywoki *et al.* 2011). The Clinician collected the data during the participant attendance at the study clinic. Additional data on anthropometric measures (weight, Mid Upper Arm Circumference (MUAC), height), oxygen saturation (using pulse oximeter), heart rate, any laboratory tests done, diagnosis and treatment given were also recorded.

iii) Household Risk Survey Questionnaire

The questionnaire was also modified from a previous study in our site (Okiro 2007; Okiro *et al.* 2008). Data on potential risk factors of virus transmission and infection in the household were collected. This included details on who and how the study infant was taken care of in the household, ownership of property and quality of the household head's house for assessing socio-economic status, presence of a toilet and waste management as well as source of water for domestic use. Individual level characteristics, such as whether in school, smokes, and anthropometric measures such as weight, mid-upper arm circumference (MUAC) and height were recorded.

3.9.2 Data handling and entry

The project manager assisted by the study coordinator reviewed the forms at the field office. Corrective action was taken if any anomalies were identified at this stage mainly involving a revisit of the household. The filled forms were forwarded to KWTRP in Kilifi on daily basis for data entry. All the field and laboratory data were doubled-entered on a Filemaker database specially designed and coded by the project manager, in consultation with the Centre's

programmers (FileMaker Pro version 9, FileMaker Inc, US). In addition to the strict data checks at entry (e.g. set date formats, decimals allowed and provision of drop down menus etc.), random checks were regularly conducted on the database to ensure the data were accurate and up-to-date. The regular checks involved selecting randomly ~10 data collection forms every week and crosschecking against the entered data. Appendix O shows screen shots of the database. All the source documents were sorted by household identity and chronologically stored in a cabinet accessible only to the study team.

3.9.3 Data cleaning and analysis

The double entered data were exported in comma separated values files and loaded into STATA (version 11.2, STATACORP, College Station, Texas, US) for data cleaning. Original forms were used to resolve any disparities in the two entries. The clean data with the household, participant, visit, sample level variables and laboratory results were used for subsequent analysis. Table 3.1 in this Chapter shows the definition of terms that formed the guide to structuring of the data in readiness for analyses. The specific analysis plans are presented in the methods section of the subsequent Chapters. All analyses were done in STATA version 11.2 unless otherwise stated.

In this chapter, we analysed data arising from samples collected in a study conducted in Pingilikani Health Centre, an outpatient clinic, where the original objective was to assess the diagnostic performance of NPS relative to NW (Munywoki *et al.* 2011). In the outpatient study, a total of 299-paired samples were collected, RNA extracted using Qiagen kit and tested using the M-PCR assay. For the current work, we selected 30 archived nasal samples at random and extracted RNA using three different kits (QIAamp Viral RNA Mini (Qiagen) Kit, MagNa pure LC Total nucleic acid (TNA) and MagNa pure LC High performance (HP) kits). This was to assess the sensitivity of the various kits in detection of RSV and other

respiratory viruses. Another set of 112 nasal specimen were also randomly selected and RNA extracted using the HP kit and divided into two aliquots: one was used for uniplex and the other for triplex real time RT-PCR to test for RSV A, RSV B and adenoviruses. This second set of samples was made to assess the effect of multiplexing in detecting of RSV and other respiratory viruses. Specimens were assigned positive for a particular pathogen if the Ct value was ≤35.0. A sample was considered a true positive if either of the extraction methods was positive and comparisons made using McNemar's chi-square test. The Binomial Exact method was used to determine 95% confidence intervals (CI) for the sensitivities (one-sided 97.5% reported if sensitivity was 100%). The mean (95% CI) of the Ct values by extraction or the screening method was calculated and comparisons made using paired t-test for each virus.

3.10 Ethical considerations

The consenting process was thorough and participants were allowed to seek clarification or withdraw at any stage during the study. The key areas considered in planning and conduct of the study are highlighted.

3.10.1 Risks

No procedures used in this study represented a significant risk to the participants. The nasal sampling could result in mild discomfort which was explained to each participant prior to collection. Suitably trained staff performed all study procedures, and these procedures were carefully explained in advance to the older children and adults. The frequency of home visits was discussed with the household head and all individual members so that a convenient schedule was agreed upon. The scheduling was flexible and would vary from week-to-week depending on participant's availability.

3.10.2 Benefits

There were no direct benefits accruing from the frequent nasal/oral sampling and testing since there is no established treatment for viral respiratory infections. However, the participants may well have benefited from close monitoring for any illness by the study coordinator (a nurse) during the home visits and were referred to the study clinic promptly for treatment when necessary. All costs of treatment including drugs for acute illnesses incurred in Matsangoni dispensary were paid for by the project. Expenses for referral to KDH and costs of inpatient treatment were also incurred by the project. The community, in general, benefited from enhanced supply of drugs to the dispensary. The project supported the Ministry of Public Health and Sanitation in dealing with a cholera outbreak that occurred in the area during the study period. We provided vehicle for ease of referral of severe 'suspected' cholera cases to KDH and more substantially provided supplies used in cleaning and covering of open water wells. Other benefits to the study households were in form of reading books and pens to the school going children but our main efforts to promote education in the area was through supply of text books and holding of career days in each of the six public primary and one secondary school in the area. In line with local culture we recognised the household involvement by offering a gift hamper at the end of the study. This was not planned for at the outset and the CAST (while planning on the household and community exit strategy) decided on the package. This consisted of household consumables that would last an average family (8 members) about 1-2 days.

3.10.3 Data confidentiality

Data were stored on password-protected database, accessible only to study investigators and under close monitoring by the project manager. Data were archived by the Centre's database administrator. Project manager created the analysis databases with all personal identifiers

removed from the main study database. All the source documents were stored in cabinets under lock and key.

3.10.4 Feedback of information

The individual test results of the nasal and oral fluid samples were not conveyed to the participants for they were not likely to affect their health care. This was clearly explained during the consenting process. A feedback meeting was organised at the end of the study to share the major study experiences and results.

3.10.5 Scientific and ethical approval

The following scientific and ethical committees approved the project:

a) The *Scientific Coordinating Committee (SCC)* based in KWTRP, Centre for Geographic Medicine and Research - Coast (CGMRC) in Kilifi, Kenya, reviewed and approved the study proposal on 5th June 2009. The affiliated Communications and Consent Committee (CCC) had reviewed the ethical aspects of the study and recommended for approval prior to the SCC meeting.

b) *KEMRI Scientific Steering Committee (SSC),* based in KEMRI headquarters in Nairobi, Kenya, approved the scientific basis of the study on 30th June 2009. The registration number for the study is: SSC no. 1651.

c) *National Ethical Review Committee (NERC),* which is an independent Kenyan ethics board, approved the study on 16th September 2009. Ethical approval has been since renewed twice in accordance to the local ethics guidelines. The current approval has been granted until 12th August 2013.

d) *Biomedical Research Ethical Sub-Committee,* University of Warwick, UK. This committee reviewed and approved the study on 8th September 2009.

All the study approval letters are provided in Appendix P

3.10.6 Consent forms and consenting process

Informed consent documents used in this study (Appendix G) were developed and submitted for approval to the consent review committees (see the above committees) as an appendix to the proposal. The consent forms were later translated from English to Kiswahili and Kigiriama, which are languages used by the local community. The process of translation and back translation was done jointly by the candidate and the experienced[4] field workers, who are from the local community and are fluent in both languages. The household study field workers (as they were commonly referred) also proof read and made the final corrections in consultation with the project manager and the final documents reviewed by the CLG, whose members form part of Centre's Communication and Consent Committee. This final review was mainly to ensure the original message in the English consent forms was preserved during the translation process.

Informed consent was obtained from the household head and all members of the household only after adequate explanation of the study procedures. The study was also explained to those aged less than eighteen years (as appropriate) and their verbal assent obtained before a signed informed consent was obtained from their parents or guardian. Specific consent was sought for storage and later use of the samples.

---

[4] *Field workers who have been working in KDH paediatric wards particularly those who were involved in the conduct of RSV surveillance studies since 2002*

3.11 Quality control on sample and data collection

Before inception of the study, all the field workers were trained on the conduct of the study including its rationale, objectives, procedures and the required documentations. They had instructions on filling the data collection forms at all times and the study coordinator and the project manager crosschecked the filled forms regularly. The field workers were trained to be methodical in their actions as well as to make sure they were clear on exactly what they intended to do before arriving at the household. At the end of each visit, a review of the entry was carried out before leaving the home. On the return, to the field office at Matsangoni dispensary, the field workers had to (i) go through their individual home visit forms to ensure that no obvious mistakes had been made (ii) relate any problems encountered in the field to the senior field worker or the study co-ordinator or the project manager and also note them down for discussion during the weekly meetings and, (iii) enter the details of participants from whom NPS and OF had been collected on each day into the sample reception book in the dispensary.

The senior field worker, study coordinator and the project manager reviewed the Home Visit forms and sample reception log prior to their transfer to KWTRP in Kilifi for data entry. Any errors were referred to the respective field worker for immediate corrective action. During the initiation and regularly in the course of the study, the project manager and study coordinator accompanied the field workers on the home visits to ensure that study questions were asked correctly, that samples were collected in proper manner and that the information provided by the participants was entered into the forms correctly. Any mistakes made were reviewed and the fieldworker advised accordingly. Training of the field workers on study protocol and sample collection was done periodically throughout the whole study period.

Strict measures were taken to ensure that (i) all bottles and tubes for NPS and OF specimens were labelled properly before any sample was taken, (ii) all samples had the participant's name, serial number, sample designation and the date of sample collection, and (iii) saliva samples were collected for one minute to ensure uniformity. NPS were collected once a week from all the field team members including the project manager for viral testing using M-PCR. Any infected staffs were relieved off their duties of home visits until recovery (tested negative) to minimise chances of infection transmission into the households by the field team themselves. Other measures to minimise infection transmission and cross contamination in samples included use of gloves in specimen collection and hand washing (using hand sanitizers after every specimen collection). Household members were not allowed handle the field worker kit.

Data entry was into a specially designed Filemaker database with logical and range checks and flags for the reduction of entry errors. The project manager checked the data regularly and double entry was done. The two entries were compared and any disparities corrected using the original source documents.

3.12 Challenges and how they were addressed

A number of challenges were encountered in the course of the study, most of which are highlighted below:

a) Difficulties in community entry: Though we had an elaborate strategy of community engagement and sensitisation at the outset various challenges were encountered. This included some community members spreading rumours on the nature of our work. Our inclusion criteria was challenged in certain areas because of the general concern in the local population that the Centre's research largely 'targets' families with young children.

Community members requested to see more studies involving adults who are similarly affected by diseases. Following several meetings with various community groups we were able to explain that young children experience a higher burden of diseases and are at an increased risk of severe and fatal outcomes relative to adults hence the tendency to focus research on this vulnerable age group.

b) Poor acceptability of nasal swabbing: This was one of the anticipated challenges. During the first week of sampling when six households withdrew, most citing dislike of the swabbing, it appeared this would be a real obstacle to the implementation of the study. Sensitisation on the study procedures at this point was heightened and perceived fears addressed. Field workers were regularly monitored (and retrained when necessary) to ensure the nasal swabbing was done well, i.e. specifically that the swab was inserted sufficiently deep and with minimal discomfort. We also started sample collection (once-a-week) within the field team members themselves. This helped in rationalising the anxiety, fear and discomfort reported during nasal swabbing of the participants as well as an opportunity to have peer assessment of the quality of sampling. As the field workers gained experience, sample collection was infrequently raised as a concern.

c) Lack of fridge (electricity) in Matsangoni dispensary for ice pack and sample storage. Initially it was envisaged samples would be transported to the Kilifi laboratory at the end of every day. However, this was found very difficult for the samples collected after working hours (after 5pm). Matsangoni dispensary had no electricity supply hence a gas-powered fridge was purchased for sample and ice pack storage, distributing workload over the entire working hours and beyond.

d) Demanding workload. Long working hours were required in part because of the need to fit into the working day of household members and thus visits early and late each day.

This necessitated recruitment of additional field staff and overtime payment of the field workers.

Overall, the above challenges were discussed and solutions found in our weekly meetings. A CAST representative regularly attended the weekly meetings especially during the study initiation and was invited when community engagement issues were being discussed. This was a very important process not only for review of performance and set weekly targets but to share good practice and any field issues facilitating early detections of community concerns.

3.13 Results

3.13.1 Sensitivity of M-PCR in detection of RSV by RNA extraction method

Previously in Kilifi (Munywoki *et al.* 2011), viral RNA was extracted using the Qiagen extraction method, but due to the large number of samples in the current study, we acquired a semi-automated extraction machine that uses either total nucleic acid (TNA) or high performance (HP) extraction kits. Before deciding on the kits to use, we conducted an experiment to establish the diagnostic performance of the TNA and HP extraction kits in detection of RSV A/B against the Qiagen method. Thirty nasal samples collected from an outpatient setting study (Munywoki *et al.* 2011) were randomly selected. The selection was limited to only samples previously extracted by Qiagen manual extraction method and recording a cycle threshold (Ct) value of above 0. The samples were screened using 'RSV A/B and adenovirus' triplex after separate extraction of the samples by Magna Pure LC TNA and HP extraction kits. Twenty-six samples were positive for either RSV A or B by Qiagen (Ct value of <35). Of these positive samples, 17 (65%) and 23 (88%) were found RSV positive (using the same Ct cut off value of <35) by TNA and HP extraction kits, respectively. Detailed tabulation of the results using varying cut offs are shown in Table 3.4

below. On comparison of the Ct values from the three extraction methods (Figure 3.10 and Table 3.5), the HP kit had similar mean Ct values relative to Qiagen (25.76 vs. 26.26, respectively). The TNA kit yielded higher Ct values (Mean Ct value of 31.68).

A further attempt to compare the Qiagen and HP using a large set of nasal samples (n=111) yielded similar results (see results in Appendix Q).

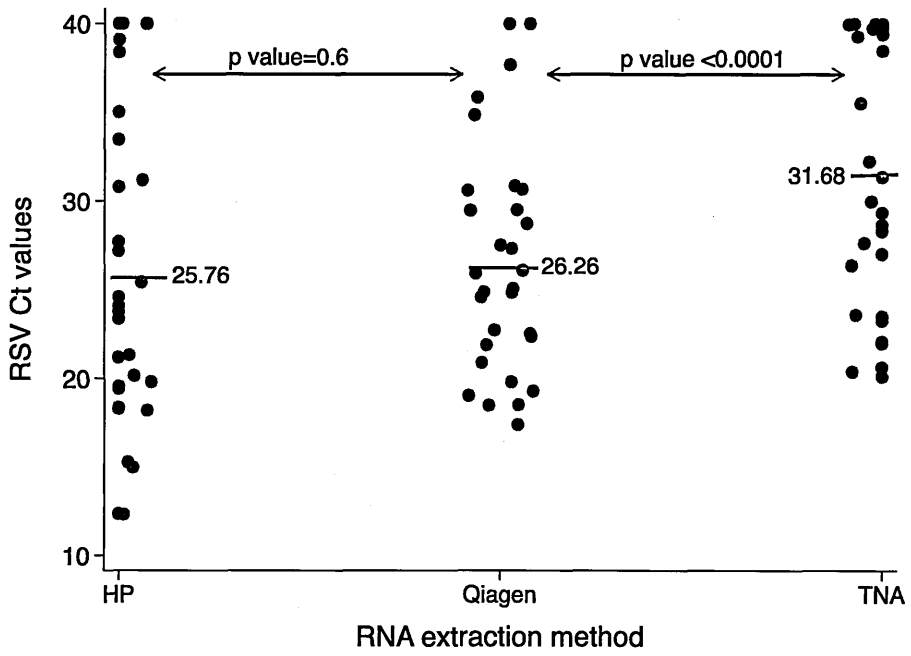


Figure 3.10: Scatter plot showing the distribution of Ct values for RSV A/B detection using RNA from the three extraction methods. The red lines show the mean Ct values. The Qiagen extraction method is used as the reference for the paired t-test.

Table 3.4: Number of virus positives (sensitivity) at varying Ct values by the extraction methods for the screened viruses

	Qiagen, n (Ref)				**TNA, n (%)**				**HP, n (%)**			
Cut-offs[1]	*<25*	*<30*	*<35*	*<40*	*<25*	*<30*	*<35*	*<40*	*<25*	*<30*	*<35*	*<40*
RSV A	9	11	12	16	2 (22)	7 (78)	9(75)	11(69)	9(100)	12(>100)	13(>100)	14(88)
RSV B	6	12	14	14	6(100)	8(67)	8(57)	8(57)	6(100)	9(75)	10(71)	10(71)
RSV	15	23	26	30	8(53)	15 (65)	17(65)	19(63)	15(100)	21(91)	23(88)	24(80)

Key: n, number of samples with RSV detections; %, show in brackets, are the sensitivities of the extraction methods in detection of RSV for TNA and HP using Qiagen RNA extraction method as the gold standard; Ref, reference; 1, the shown cut offs of the Ct values across the table are based on the results from Qiagen method; Ct, Cycle threshold

Table 3.5: Comparison of mean Ct values for detection of RSV by extraction method

	Qiagen		TNA			HP		
	Mean	*95% CI*	*Mean*	*%95 CI*	*P value*	*Mean*	*% 95 CI*	*P value*
RSV A	33.08	29.90 – 36.26	35.96	33.64 – 38.27	0.0145	31.54	27.67 – 35.41	0.1861
RSV B	33.20	30.21 – 36.18	35.72	32.95 – 38.49	0.0386	34.22	30.94 – 37.50	0.4026
RSV	26.26	23.95 – 28.60	31.68	28.84 – 34.51	<0.0001	25.76	22.35 – 29.17	0.6442

Key: All Ct values of 0 were coded as 40 for this analysis; Qiagen results were used as the gold standard for the reported comparisons; T-test P values for comparison of HP vs. TNA Ct values were RSV A=0.0002, RSV B =0.0055 and RSV <0.0001

3.13.2 Sensitivity of RSV detection using uniplex and triplex real time M-PCR in detection of RSV A, B and adenoviruses

A total of 111 samples were screened for RSV A/B and Adenoviruses using both uniplex and triplex assays. The samples were randomly selected from the Pingilikani outpatient study ensuring a similar representation of the infection status: Positive sample for RSV A (30), RSV B (29), Adenovirus (22) and the rest were negative for the three viruses based on the initial M-PCR testing (Munywoki *et al.* 2011). Most (88.3%) of the samples used in this experiment were NW. Male participants were 50 (45.1%). The median age (interquartile range) of the 102 participants with recorded birth date was 1.8 (0.8 to 3.9) years. The youngest was 1.3 months and oldest 10.6 years. Infants were 35 (34.3%).

A set of 60 samples was screened using uniplex real time PCR for RSV group A and B while 51 samples for adenoviruses uniplex. All the samples (111) were also screened by 'RSV A/B and adenoviruses' triplex real-time M-PCR. The uniplex and triplex assays had comparable sensitivities in detection of RSV A (95.0% vs. 92.5% respectively) and RSV B (97.0% vs. 100%, respectively), Table 3.6. However, the adenovirus uniplex assay was more sensitive than triplex method (100% vs. 75% respectively) and this was statistically significant (p-value= 0.0156). These results were reflected in high correlation of the Ct values for RSV group A and B (97% and 98%, respectively) unlike for adenovirus (66%) when comparing uniplex Ct values with triplex Ct values (Figure 3.11). However a detailed analysis showed Ct values for the triplex assay were consistently higher than for uniplex and that was significant for RSV group B and adenoviruses, p values =<0.0001 and 0.0003 respectively (Table 3.7). However, these differences were in the 1 – 2 Ct range.

Table 3.6: Respiratory pathogens detected and their sensitivity by screening method

Detections		**Uniplex sensitivity**		**Triplex sensitivity**		***P value*[3]**
Virus (samples[1])	*N*	*n*	*% (95% CI[1])*	*n*	*% (95% CI[2])*	
Any virus (111)	95	90	94.7 (88.1-98.3)	86	90.5 (82.8-95.6)	0.424
RSV A (60)	*40*	*38*	*95.0 (83.1-99.4)*	*37*	*92.5(79.6-98.4)*	*1*
RSV B (60)	*33*	*32*	*97.0 (84.2-99.9)*	*33*	*100 (89.4-100)*	*1*
Adenoviruses (51)	28	28	100 (88.1-100)	21	75.0 (55.1-89.3)	0.0156

Key: N, total number of virus detections by either uniplex or triplex which is used as the reference for calculation of sensitivity; n, number of positive samples by the respective method; %, sensitivity; CI, confidence interval; RSV, respiratory syncytial virus; 1, number of samples tested; 2, one-sided 97.5% CI reported sensitivity is 100%; 3, exact McNemar's significance probability values reported.

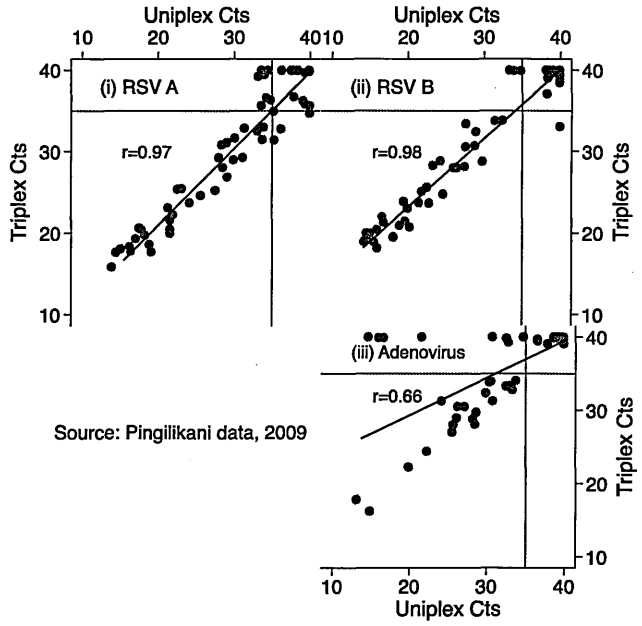


Figure 3.6: Correlation of uniplex with triplex CT values for RSV A, RSV B and adenovirus.

Table 3.7: Comparison of the Ct values by the screening method

	Uniplex	Triplex	
Virus	Mean (95% CI)	Mean (95% CI)	**P value**
RSV A	30.07 (27.86 – 32.28)	30.43 (28.29 – 32.56)	0.1555
RSV B	30.19 (27.70 – 32.68)	31.76 (29.62 – 33.89)	<0.0001
Adenoviruses	32.01 (29.70 – 34.33)	35.42 (33.64 – 37.21)	0.0003

Key: N, total number of virus detections by either uniplex or triplex used as the reference for calculation of sensitivity; n, number of positive samples by the respective method; %, sensitivity; CI, confidence interval; RSV, respiratory syncytial virus; adeno, adenovirus; 1, one-sided 97.5% CI reported sensitivity is 100%; 2, exact McNemar's significance probability values reported.

3.13.3 Comparison of Ct values for IFAT positive and negative samples

The IFAT method detects RSV virus in the infected epithelial cells signalling that the virus was replicating (and most likely infectious) while the PCR methods detect nucleic acid material (in infected epithelial cells or lying in the upper airway). Analysis of the Pingilikani data was made to assess the Ct value (from M-PCR) distribution in IFAT positive and negative samples. The 43 NPS collections positive by IFAT were PCR positive with Ct values of less than 30 except for two NPS collections that had no PCR amplification. Additional 31 samples were found RSV positive by M-PCR from the 256 NPS collections that were RSV negative by the IFAT method. Examining the Ct values further (Figure 3.12) revealed IFAT positive had lower mean Ct value compared to the IFAT negatives (22.69 versus 29.19, respectively; t-test p-value<0.0001). Older children (≥2 years) had higher median Ct value relative to the younger counterparts but the difference was not statistically significant (Table 3.8 and Figure 3.13).

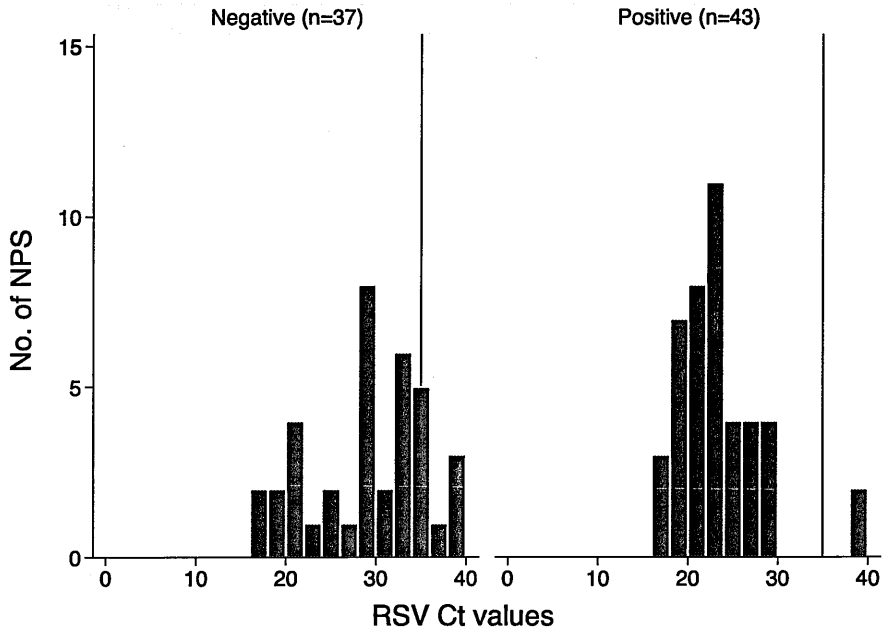


Figure 3.12: Histogram of the RSV Ct values by IFAT status. The red vertical line shows the cut off for virus positivity. IFAT negative samples with no amplification signal (n=219) were excluded from the histogram

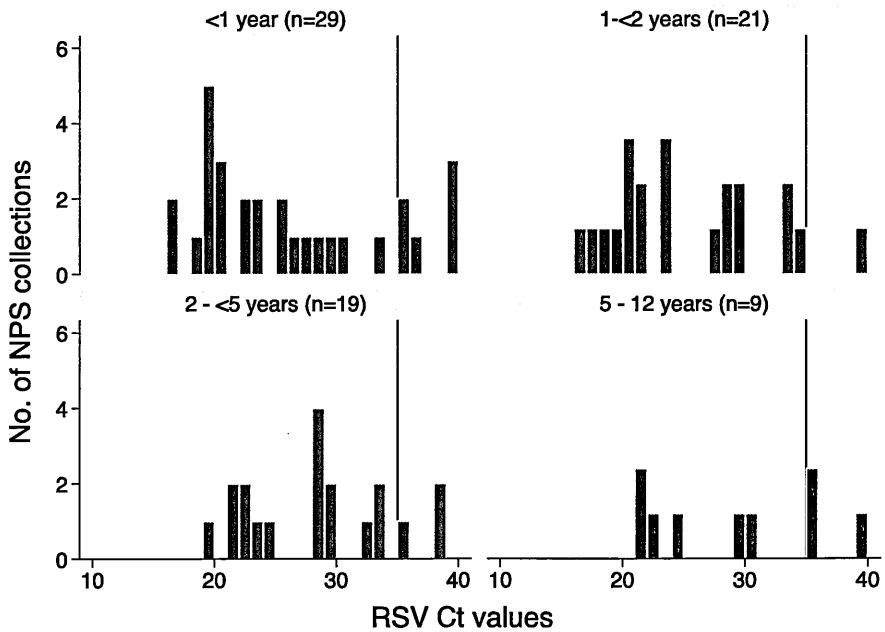


Figure 3.13: Frequency distribution of RSV Ct values by age class. The vertical red line indicates Ct value of 35, cut off for viral positivity

Table 3.8: Comparison of Ct values in 78 nasal samples stratified by age

Age classes	N	Mean[1]	Median[2]	IQR	
<1 year	29	26.14	23.74	20.00 – 30.42	Ref
1 – <2 years	21	25.41	23.46	20.73 – 29.57	0.8983
2 – <5 years	19	28.53	28.5	22.7 – 33.21	0.1429
5 – 12 years	9	28.84	29.22	22.55 – 35.4	0.2501
All	*78*	*26.84*	*25.32*	*21.00 – 32.32*	-

Key: N, number samples with a Ct value; IQR, interquanrtile range; 1, mean Ct values; 2, median Ct value

3.14 Discussion

The study called for high level of community interaction and a structured community engagement process was developed at outset and followed with the aim of involving most of the stakeholders. Eligible households were recruited before the start of the RSV season of 2009/2010 with sampling initiated immediately after the start of the epidemic which was established through the continuous inpatient RSV surveillance at Kilifi District Hospital (KDH) (Nokes *et al.* 2009). Follow-up of study participants was through active home visits by a study field team and passive participant visits at the study clinic. Sensitive and specific molecular methods were used to screen the nasal swabs for a range of common respiratory viruses including RSV. The elaborate approach in the study design and implementation involved careful community sensitization, adequate training of the field team on communication and sample collection procedures, thorough consenting process and continuous feedback from study participants, field staff and community leaders. Weekly field team meetings closely monitored the study progress and provided a forum to share experiences and

offer customised response of any concerns was raised. This careful approach yielded the high retention rates, as it will be shown in the next Chapter.

The pilot phase of the study allowed time to set up the logistics for frequent home visits and sample collection. However, the active surveillance was not easy in study phase as scheduled home visits were not always successful and repeat visits were often arranged which called for more time. This extra load necessitated recruitment of additional field workers in the course of the study. Even though we had one vehicle and one motorbike available on weekdays for the project work, it is the use of bicycles for follows up (due to poor road network) that facilitated access to all households. This also proved to be quite cost-effective.

The quality of samples was important for this study and standardised methods were used. The field workers had a month in hospital settings for practice. The pilot phase also helped in acquitting themselves with field conditions. The integrity of the samples was assured by cold chain storage to maintain adequate epithelial cells: though not major issue when using PCR methods. Despite the lack of electricity supply in the study clinic, a 'gas powered fridge' for temporarily storage was used. Samples were transported to the main laboratories at KWTRP in Kilifi town within 24 hours of collection.

The set frequency of sampling in order to identify most infection cases, and help in identifying from whom the infections, arose from careful consideration of the reported RSV shedding durations (Okiro *et al.* 2010). However, infection data would be missing in some instances, as the sample collection at all participants' contacts was not possible. In addition, acceptability of NPS collection method among the adults and teenagers was initially challenging but our patience in explaining the procedure

seemed to have helped to improve compliance. It is also not clear to what extend our decision to offer textbooks to the school going and sweets to the young children during the home visits influenced the acceptability rates. We have shown in our setting that the NPS is the preferred (~70%) method by both parents/guardians and the children relative to NW collection method (Munywoki *et al.* 2011).

We used sensitive molecular methods in detection of viruses. Earlier studies used culture (Hall *et al.* 1976) and more recent work involved use of IFAT (Okiro 2007). These methods (culture and IFAT) have been shown to be less sensitive to PCR (Falsey *et al.* 2003) and more likely to miss infections associated with low viral load particularly in older children and adults (Stockton *et al.* 1998; West *et al.* 1998; van Elden *et al.* 2003). Molecular techniques detect viral nucleic acid and thus the increased sensitivity of the PCR raised questions about the viability of the detected virus in causing the infection. This is an area that would require further investigation but data from our inpatient studies suggest IFAT identifies participants with high viral load and PCR has additional benefit of even detecting individuals with low viral titre as depicted by distribution of Ct values for IFAT positive vis-a-vis IFAT negative samples. The M-PCR method is expensive but the study was adequately funded by the Welcome Trust. Further screening of the nasal samples from the Pingilikani study, provided invaluable information on the diagnostic performance of the adopted M-PCR assay. We showed that both the uniplex and triplex had good sensitivities (above 90%) in detecting RSV with no statistical difference in sensitivity of detecting RSV A and RSV B by either method. However, there was a slight increase in the sensitivity of uniplex in the detection of adenoviruses compared with triplex ($p=0.0156$). Though uniplex seems superior in producing better (lower) Ct values in detection of RSV B and adenovirus, the method was labour intensive and costly and therefore lacks merit

in implementation in a large epidemiological surveillance study. The triplex thus, offered a suitable alternative for detection of RSV A and B but further evaluation of the adenovirus detection is warranted. On extractions, the HP kit was more sensitive in detection of RSV relative to Qiagen and TNA kits. Overall, the HP kit had better (lower) Ct values compared to Qiagen kit for detection of RSV A, RSV B and adenovirus. Therefore these data favoured adoption of HP extraction method and indicated non-inferiority in use of our routine real time M-PCR method in detection of RSV A, B and adenovirus.

The candidate has experience in field work and sample collection both in hospital and community set up (Nokes *et al.* 2009; Berkley *et al.* 2010; Munywoki *et al.* 2011; Munywoki *et al.* 2013). This helped in designing and conducting staff training and in the initiation and monitoring of the study progress with particular focus on ensuring the study protocol was observed throughout. For instance, experience in working in hospital setting helped in training on recognition of respiratory symptoms and sample collection, thus standardising the study procedures. Moreover, data collection was with a standardized proformas, which have been in use in our site, with questions adapted to suit the community study and piloted for clarity and suitability. A manual was used as the day-to-day reference guide for the field team.

In conclusion the study employed an elaborate study procedures and the choice of real time multiplex PCR would offer advantage in increased sensitivity to detect the target viruses.

CHAPTER FOUR

4 General Results

4.1 Introduction

Household studies are dependent on assessment of infection (s). Increased sensitivity and range of pathogens detectable by molecular diagnostics over traditional method has boosted the study of virus epidemiology. In order to identify infections in the household and the possible chains of transmission, detailed surveillance is required. Such surveillance work requires use of suitable and acceptable sample collection methods and collection of samples regardless of illness. Reliance on the occurrence of symptoms to identify infected persons might result to failure in recognizing sub-clinical infections and to identify links in transmission events. However, there are very few studies or respiratory viruses which have collected samples frequently and irrespective of symptoms (Hall *et al.* 1976). As a consequence, estimates of the proportion of RSV infections that are subclinical are limited, and transmission patterns in households are poorly mapped out, especially in older children and adults where shedding levels are less than for younger aged individuals. Thus, these issues find focus in this study. Even with the intensive sample collections, it is possible that some infections, especially mild infections associated with short duration of shedding (less than 3 days) or sub-clinical in older children and adults with low viral titers are likely to be missed. Efforts to capture these infections require alternate approach such as measuring antibody responses in serum. However, blood collection is invasive and regular sampling would not be acceptable in the community. Oral fluid (OF) has been suggested as alternative specimen for RSV-specific antibody measurement (Okiro *et al.* 2008) and even for viral detection using molecular methods (von Linstow *et al.*

2006) and was collected once a-week in the current study to supplement infection data. The diagnostic performance of this painless collection method was also investigated.

This chapter provides the general results of the household study, which lay the foundation for the findings presented in the subsequent Chapters.

4.2 Chapter outline

In this Chapter, baseline characteristics of households and the study participants are reported. Temporal distributions of RSV and other respiratory viruses are shown with particular focus on co-infection with RSV. Included also are the general study results related to RSV infections, which set the scene for the findings presented in the subsequent Chapters. The proportion of subclinical infections by age and sensitivity of oral fluid in the detection of RSV by M-PCR relative to NPS is also presented.

4.3 Methods

Detailed descriptions of the study methods are provided in Chapter 3.

4.3.1 Data analysis

STATA Version 11.2 (StataCorp, College Station, Texas, USA) was used for all data analyses in this Chapter. Student's t-test, Wilcoxon rank-sum test, chi-square test, and Fisher's exact test were used as appropriate. The terms and their definitions used in this analysis were described in Chapter 3. Additional terms specific to RSV infections are provided in Table 4.1.

Household members were categorised based on the relationship to study infants (which was the youngest child at the time of household recruitment) into six groups; namely, (i) the study infant, in some instances referred to as self, (ii) siblings, (iii) cousins and other relatives (nephew and niece) aged <15 years referred generally as

cousins, (iii) mother, (iv) father or (v) other adults in the household (which included aunts, uncles, grandparents, step-mother and other relatives aged ≥15 years). Trends in household recruitment and loss of follow up are presented in the previous Chapter. The characteristics of the retained households were compared to those lost to follow up. The NPS compliance rates were calculated as the percentage of the collected samples divided by the total expected samples based on the specimen collection regime (twice-a-week for NPS and once-a-week for OF). The frequency distributions of the Ct values for the various targets in the M-PCR assay were also examined. The Ct values distributions by age were plotted for RSV, adenoviruses, human coronaviruses (OC43, NL63 and 229E) and rhinovirus. Weekly data on virus detections were plotted to show the temporal distributions and co-circulation. Prevalence of RSV infections and other respiratory viruses in households, individuals and samples are shown. In order to calculate the risk of viral infection by clinical status at the time of sampling, NPS collections were tested for the full range of targets in the M-PCR assay. The prevalence of the detected viruses by clinical status is presented.

Primary and co-primary cases of RSV were identified as defined in Table 4.1. Any other RSV infection episodes were considered as secondary cases (also defined in Table 4.1). Household and age-specific crude attack rates were determined by dividing the number of RSV individual episodes by the total number of individuals in the respective category. Similarly, secondary attack rates constituted the number of secondary cases divided by total uninfected household contacts. Proportions of subclinical RSV infections by age and other characteristics were also calculated.

Table 4.1: Definition of terms related to RSV positivity in the households

Term	Definitions
RSV	Refers to PCR positive for either RSV groups A or B
Household episode	Period within which one or more individual episodes occurred in members of the same household without an interval of 14 days or more in which a positive specimen was absent from the household.
Household outbreak	Where more than one individual episode occurs within a household episode (i.e. where a primary infection spreads to at least one other household member)
Primary or co-primary case	The first individual episode within a household based on the dates of sample collection. If individual episodes started on the same date in ≥2 members of the same household, they were referred as co-primary.
Secondary case	Any RSV infection episodes occurring after the introduction of the virus into the household i.e. all non-primary cases.
Household contacts	Household members who were present and not infected during a household episode.

Key: RSV, respiratory syncytial virus; PCR, polymerase chain reaction

4.4 Results

4.4.1 Households recruitment

Household recruitment started on 4th November 2009 in Uyombo sub-location. Recruitment was extended to include households from Matsangoni sub-location later. Regular specimen sampling began on 8th December 2009, after consenting 50 households and establishing the RSV season had began (see Table 3.1 for definition) based on data from our inpatient surveillance at KDH. Nine households withdrew during the pilot phase, with 2 replacements. All the study procedures, including NPS

collections every 3 – 4 days, were fully implemented by 11th January 2010, the official study start date. Subsequently, a further 4 households were lost to follow-up and 8 household replacements recruited with the last recruitment taking place on the 5th March 2010 and last withdrawal occurring on 7th March 2010. The study closed on 4th June 2010 after 24 weeks of follow up, Figure 3.10. Overall, 60 households (596 participants) were recruited and 13 households (103 participants) were lost to follow-up. The distribution of the consented households by village and sub-location are shown in Table 4.2.

Table 4.2: Distribution of recruited households and participants by residency

Sub-location	**Village**	**HH[1]**	**PP[2]**	**Mean[3]**	**Median (IQR)[4]**	**Min**	**Max**
Matsangoni	Madeteni	17	131	7.7	7.0 (5 – 8)	4	28
	Sidzeni	5	51	10.2	9.0 (8 – 11)	8	15
	Kilodi	1	16	16	16	-	-
	Gongoni	1	7	7	7	-	-
	Sub-total	*24*	*205*	*8.5*	*7.5 (5 – 9)*	*4*	*28*
Uyombo	Chambuko	14	136	9.7	9.5 (7 – 11)	4	23
	Wireless	7	95	13.6	8.0 (5 – 19)	5	37
	Madeteni	7	72	10.3	9.0 (6 – 16)	4	16
	Uyombo	4	46	11.5	10.5 (9 – 14)	9	15
	Maweni	4	42	10.5	8.0 (8 – 14)	7	19
	Sub-total	*36*	*391*	*10.9*	*9.0 (7 – 13)*	*4*	*37*
Overall		*60*	*596*	*9.9*	*8.0 (6 – 11)*	*4*	*37*

Key: 1, number of households (HH); 2, number of study participants (PP); 3, the mean household occupancy; 4, median household occupancy; Min, minimum number of household members per household; Max, maximum number of household members per household

4.4.2 Withdrawn households

Of the 13 households lost to follow up, six households were never sampled (declined the initial sampling (3) or became ineligible due to separation of parents (2) or out-migrated (1)), another six households withdrew in the course of the study with the members citing dislike or fear of the frequent nasopharyngeal swabbing, and one out-migrated after sampling had started. Most (10/13, 76.9%) of the withdrawn households were from Uyombo sub-location where household recruitment started.

Comparisons of the characteristics of the retained and withdrawn households are presented in Table 4.3. Most (10/13, 76.9%) of the withdrawn households were from Uyombo sub-location where household recruitment started. The proportion of male participants in the withdrawn households was higher (57.1%) compared to the proportion in retained households, (43.8%), chi square p-value=0.002. The withdrawn households had few school-going children relative to the retained households ($p<0.001$). In all the other recorded characteristics the withdrawn (13) households were similar to the retained (47) households, Table 4.3. All the subsequent analyses in this chapter exclude data from the 13 withdrawn households. The retained households were followed for a median (interquartile range, IQR) duration of 24.6 (23.6 – 25.3) weeks.

4.4.3 Baseline household characteristics

The mean household occupancy was 10.5 members with a median (interquartile range, IQR) of 8 (6 – 12) members, Table 4. 3. The smallest households had 4 members (e.g. the study infant, a sibling, mother and father) while the largest household had 37 members. The frequency distributions of the household sizes are shown in Figure 4.1. The households comprised of one or more related nuclear families (parents and children units) as shown in Table 4.4. The prevalence of males

per household had a mean of 43.8% (95% CI, 36.4 – 57.1%) while the median (IQR) number of older children per household was 4 (3 – 6). On average, the household members in each household were 15.5 (95% CI, 13.2 – 17.9) years old. The youngest participant was 13 days old and the oldest 92 years, at recruitment. About a quarter (12/46) of the mothers to the study infant had no formal education i.e. had never attended school. Uyombo sub-location had bigger household sizes compared to Matsangoni, median (IQR) of 9.0 (7 – 13) and 7.5 (5 – 9), respectively, although the difference was not statistically significant ($p = 0.1475$).

Table 4.3 Baseline characteristics of the recruited households

Characteristic[1]	In study (N=47)	Lost to follow up (N=13)	P-value
Households from Uyombo, n (%)	24 (51.1)	10 (76.9)	0.087
Household sizes	8 (6 – 12)	8 (6 – 9)	0.276[2]
Children (<15y) per HH, mean (95% CI)	6.4 (5.3 – 7.5)	4.9 (3.6 – 6.3)	0.172
Children (1– 14 years) per HH	4 (3 – 6)	4 (3 – 5)	0.204
School-going children per HH	4 (3 – 6)	0 (0 – 1)*	<0.001[3]
Male: female ratio per HH	0.78 (0.57 – 1.33)	1.33 (1.25 – 2.00)	0.019
Male members per HH (%)	43.8 (36.4 – 57.1)	57.1 (55.6 – 66.7)	0.018
Average age in years of members in each HH, mean (95% CI)	15.48 (13.2 – 17.9	15.3 (11.8 – 16.6)	0.597
Mothers with no formal education, n (%)	12/46[4] (26.1)	3/13 (23.1%)	0.659

Key: 1, median (IQR) statistic reported unless otherwise stated; 2, P-value based on Wilcoxon rank sum test; 3, Fisher's exact test; 4, in one household the study infant was staying with her grandmother and not included in the denominator

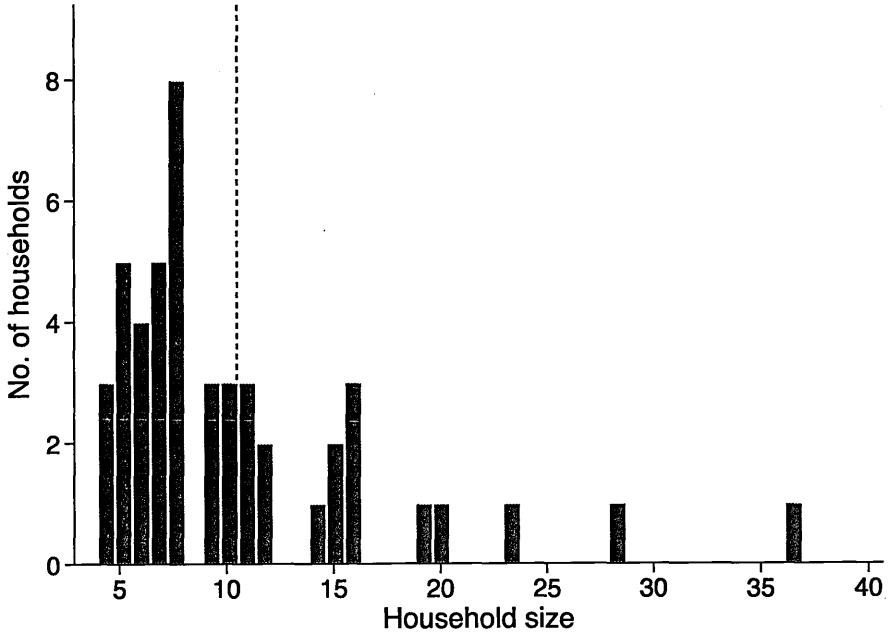


Figure 4.1: Frequency distribution of the size of the 47 retained households. The red-dotted line denotes the mean household size

Table 4.4: Distribution of the number of families per household in the 47 retained households by size

Household sizes	**No. of nuclear families per household**				
	1	2	3	4	5
4	3	0	0	0	0
5	5	0	0	0	0
6	5	0	0	0	0
7	4	0	0	0	0
8	8	0	0	0	0
9	2	1	0	0	0
10	2	1	0	0	0
11	2	1	0	0	0
12	2	0	0	0	0
14	1	0	0	0	0
15	0	2	0	0	0
16	2	1	0	0	0
19	1	0	0	0	0
20	0	0	0	1	0
23	0	0	0	0	1
28	0	0	0	0	1
37	0	1	0	0	0
All households	**36**	**8**	0	**1**	**2**

Key: HH, household; n, number of households

4.4.4 Baseline characteristics of the study participants

The 47 retained households had a total of 493 participants. The age distribution of the cohort is shown in Figure 4.2. The characteristics of the study infants are presented first followed by the other household members.

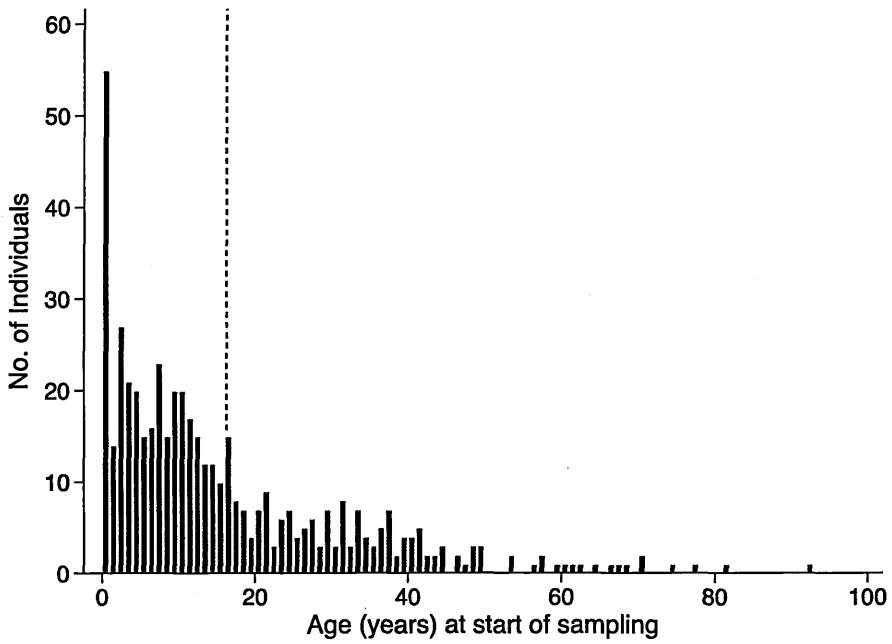


Figure 4.2: Age distribution of the 493 study participants. The red line denotes the mean age (see text for details)

i) The study infants

The study infant was the youngest child from each household at the time of recruitment. The median (IQR) age at recruitment was 3.9 (2.3 – 6.4) months. The age ranged from 13 days to almost 10 months, Figure 4.3(a). Twenty-two (46.8%) were boys. All the study infants were born after 1st April 2009 (see Figure 4.3(b)). The 2008/09 RSV epidemic ended in the first week of April 2009 and it was assumed that these children had no RSV exposure before the start of the study (as presented in Chapter 3). In 42 of the households, the study infant was the only child aged less than one year, while four households had one additional infant and

one household had four other infants. Six of the additional infants were older than the study infant while two from one household were born and recruited during the course of the study. Overall, the study had 55 children aged less than one year.

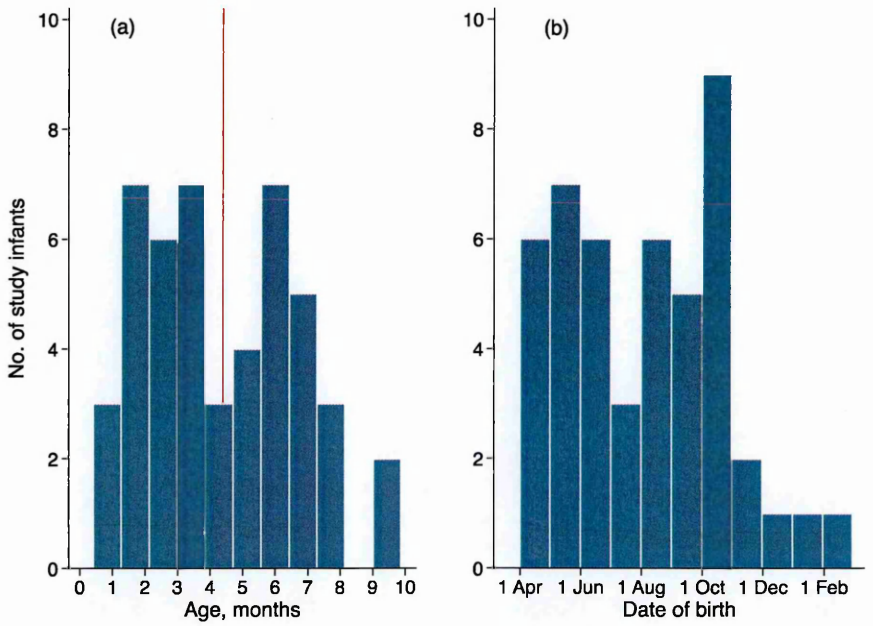


Figure 4.3: Age, at recruitment, (a) and birthdate (b) distribution of the study infants. The red line shows the mean age. The date of birth spans from 1st April 2009 to 28th February 2010

ii) Other household members

In addition to the 47 study infants, 446 other household members were recruited. This group comprised siblings (164, 36.8%), cousins (124, 27.8%), mothers (46, 10.3%), fathers (33, 7.4%) and other adult relatives (79,17.7%) of the study infants. The median age of these household members, at recruitment, was 12.6 (IQR, 6.3 – 25.6) years, and the average age was 17.9 years. Figure 4.4 shows the age distribution stratified by household relationships. This group comprised 199 (44.6%) males and 171 (38.38%) were attending school at recruitment. An additional 33 children joined school in 2010. The percentage of individuals

sharing a house, sleeping room and bed with the study infant was 53.1%, 33.9% and 22.3%, respectively. More details on the characteristics of the study participants are shown in Table 4.5.

Table 4.5: Baseline characteristics of the study infants and other household members

a) Study infants	**In study**	**Withdrawn**	**P-value**
Number	47	13	-
No. of boys (%)	22 (46.8)	7 (53.9)	0.445
Age in months at recruitment, median (IQR)	3.9 (2.3 – 6.4)	5.0 (2.5 – 5.7)	0.929
Mother has no formal education	12/46 (26.1)	3/13(23.1)	0.569
Father's occupation status			
Employed	7/33 (21.2)	5/11 (45.5)	0.577
Self-employed	16/33 (48.5)	4/11 (36.4)	
Farmer	9/33 (27.3)	4/11 (18.2)	
Not employed	1/33 (3.0)	0 (0)	
b) Other household members			
Number	446	90	
Male gender, n (%)	199 (44.6)	55 (61.1)	0.003
Age in years at recruitment, mean (95 % CI)	17.9 (16.4 – 19.4)	27.9 (5.7 – 50.1)	0.059
Age groups, n (%)			
<1y	8 (1.8)	0	0.442
1-4y	82 (18.4)	21 (23.3)	
5-14y	165 (37.0)	30 (33.3)	

15-39y	147 (33.0)	33 (36.7)	
≥40	44 (9.9)	6 (6.7)	
Relation to the study infant			
Siblings	164 (36.8)	34 (37.8)	0.169
Cousins	124 (27.8)	18 (20.0)	
Mothers	46 (10.3)	13 (14.4)	
Fathers	33 (7.4)	11 (12.2)	
Others	79 (17.7)	14 (15.6)	
Occupation status			
Unemployed	164 (36.7)	40(44.4)	0.305
Employed	16 (3.6)	5 (5.6)	
Self-employed	34 (7.6)	6 (6.7)	
Peasant farmers	61 (13.7)	10 (11.1)	
In school[1]	171 (38.3)	29 (32.2)	
Highest education level			
None	171 (38.3)	35 (39.9)	0.024
Primary	261 (58.5)	53(58.9)	
Secondary	14 (3.1)	2 (2.2)	
Living arrangement[2]			
In same house	237 (53.1)	55/88 (62.5)	0.107
Sleep in same room	151/445 (33.9)	47/85 (55.3)	<0.001
Sleep in same bed	99/445 (22.3)	34/85 (40.0)	0.001

Key: 1, additional 33 children joined school in 2010; 2, judged in relation to the study infant

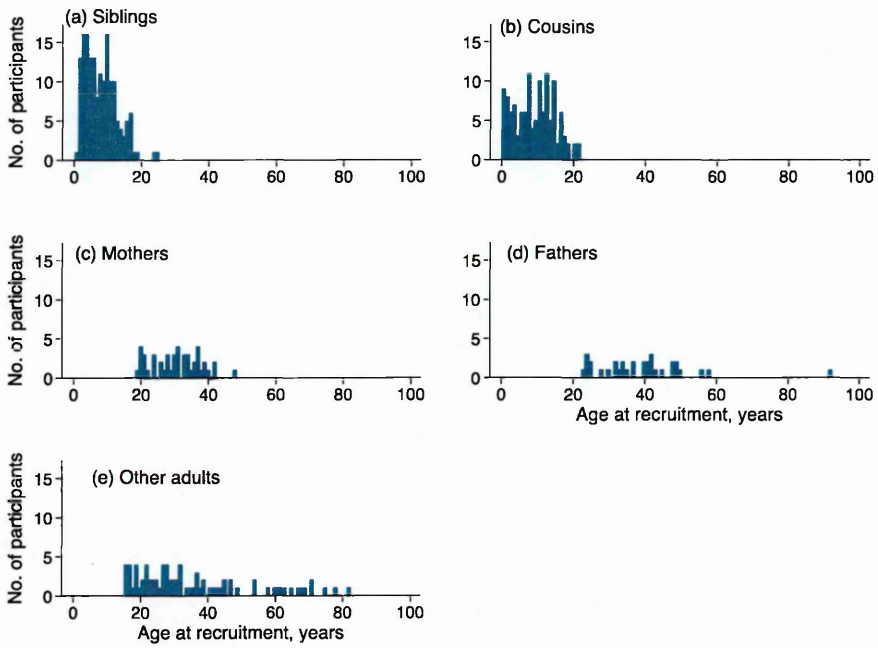


Figure 4.4: Age distributions of the siblings (a), cousins (b), mothers (c), fathers (d) and other adults (e) in the 47 study households

4.4.5 RSV epidemic of 2009/2010

The start and end of the RSV season was based on data from a RSV surveillance study in paediatric wards of KDH. Matsangoni is one of the catchment areas of the KDH (Figure 4.5). The RSV epidemic of 2008/2009 ended in the week beginning on 6th April 2009 while the RSV epidemic of 2009/2010 began on 23rd November 2009 and ended on 4th June 2010 based on our inpatient data. These dates informed the start of pilot phase as well as the end of the current study, as shown in Figure 4.6.

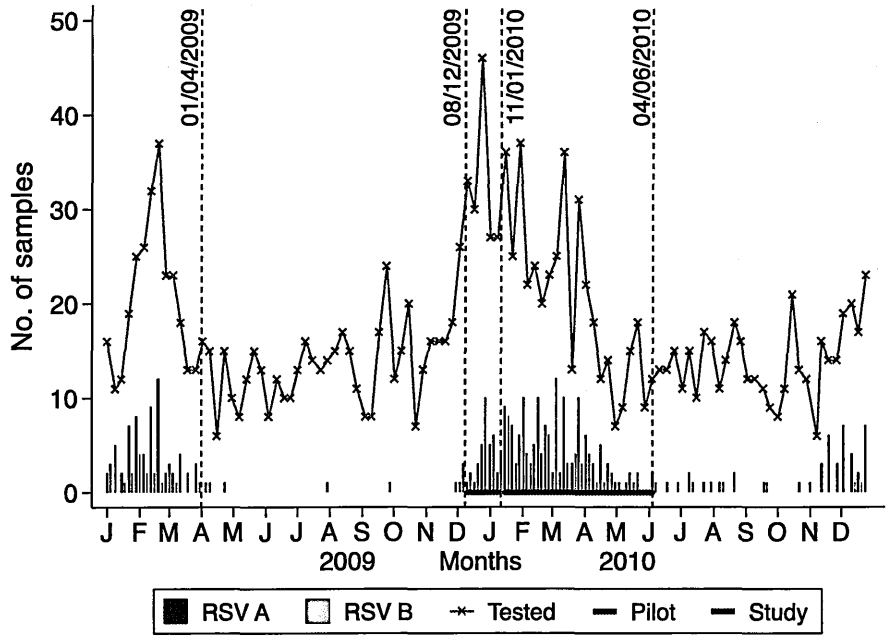


Figure 4.5: Surveillance of RSV infection in KDH paediatric wards from 1st January 2009 to 31st December 2010. The green line (with x markers) shows the weekly number of samples tested for RSV while the black and red bars shows number of RSV A and B positive samples, respectively. The green and blue bars indicate the pilot and main household study phases, respectively

4.4.6 The pilot phase

Sample collection started on 8th December 2009, when RSV season was deemed to have begun based on the inpatient surveillance data presented above (Figure 4.5). During the pilot phase, households were visited once-a-week for NPS and OF collection for a period of one month, up to 10th January 2010. A total of 1,314 individual visits were recorded resulting in 1224 (93.2%) and 1205 (91.7) collection of NPS and OF, respectively. RSV group A was detected in one NPS sample collected on 4th January 2010 from a secondary school-going individual, aged 22.0 years, (Figure 4.6).

4.4.7 The study phase

This phase started from 11th January 2010 to 4th June 2010, when RSV was prevalent in the study population, with home visits and NPS collections every 3-4 days. This period constitutes the main study data presented.

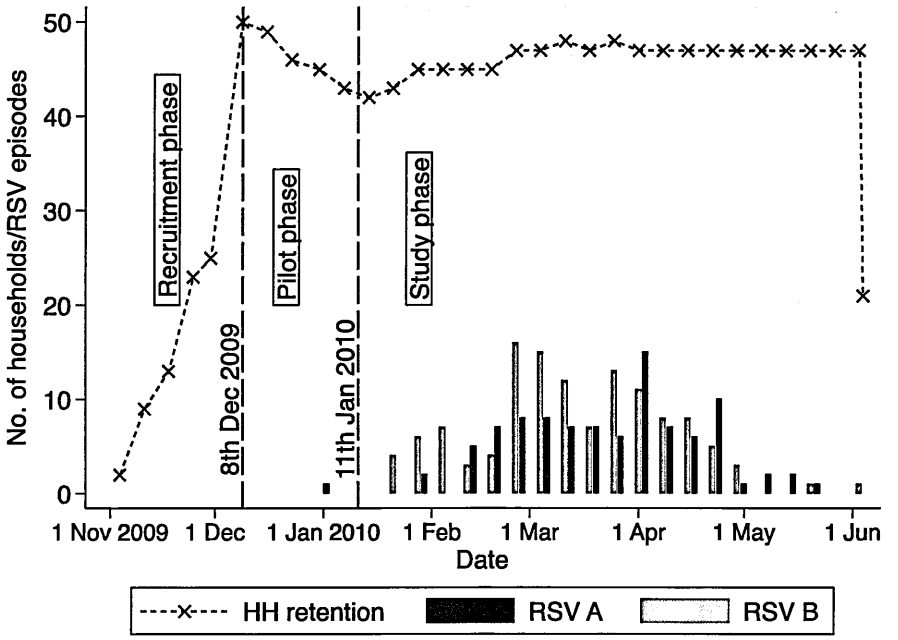


Figure 4.6: Standing number of households (HH) participating (blue dotted line) and weekly number of RSV group A (black bars) and B (red bars) individual episodes detected, over the study period

4.4.8 Home visits and sample collections

Based on our sampling regime of twice-a-week, a total of 19,707 NPS collections were expected in the study phase. In practice there was a total of 17,604 (89.3%) participant contacts and from these 16,498 (83.7 % of the expected) were collected, Table 4.6. NPS collection compliance was highest in study infants (93.5%) and lowest in other adult relatives in the households (65.1%). Fathers had 876 (67.1% of the expected) NPS collected compared to 1,716 (93.3%) from mothers. In 3,963 (22.5%) of the home visits the study participants

had one or more symptoms of ARI while in 2,304 (11.7%) of the planned visits the participants were away and neither the samples nor the health status data were collected. The median (IQR) number of home visits and NPS per individual was 42 (39 – 43) and 39 (30 – 42), respectively (Figure 4.7). Twenty (4.1%) of the participants from 10 different households had no NPS collected during the study phase even though they had one or more samples collected during the pilot phase. The trends of the number of home visits, NPS and OF collected and ARI over the study period are shown in Figure 4.8.

Table 4.6: Number of participants, home visits, NPS, OF and ARI over the study phase stratified by relationship to the study infant

		No. of participant contacts:								No. of samples:					
		Home		Clinic		With ARI		Member away		NPS collections		NPS tested[3]		OF collections	
Relation[1]	N[2]	*N*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Self (47)	1875	1787	95.3	93	5.0	918	49.0	174	9.3	1786	95.3	1686	94.4	915	97.6
Siblings (164)	6511	6284	96.5	85	1.3	1733	26.6	432	6.6	6036	92.7	5789	95.9	3139	96.4
Cousins (124)	4981	4332	87.0	59	1.2	894	17.9	632	12.7	4005	80.4	3854	96.2	2166	87.0
Mothers (46)	1839	1731	94.1	46	2.5	164	8.9	171	9.3	1716	93.3	1640	95.6	895	97.3
Fathers (33)	1306	1056	80.9	8	0.6	52	4.0	249	19.1	876	67.1	808	92.2	517	79.2
Others (79)	3195	2414	75.6	49	1.5	202	6.3	646	20.2	2079	65.1	1974	94.9	1209	75.7
All (493)	***19707***	***17604***	***89.3***	***340***	***1.7***	***3963***	***20.1***	***2304***	***11.7***	***16498***	***83.7***	***15751***	***95.5***	***8841***	***89.7***

Key: 1, relationship of the household members to the study infant. The values in brackets are the number of individual by relationship status; 2, expected number of NPS sample based on twice-a-week regime during the study phase; %, percentage was calculated from the number of NPS, OF, or ARI divided by expected number of NPS. Note the OF samples were collected once-a-week. For the NPS tested the percentage was based the number of NPS collected

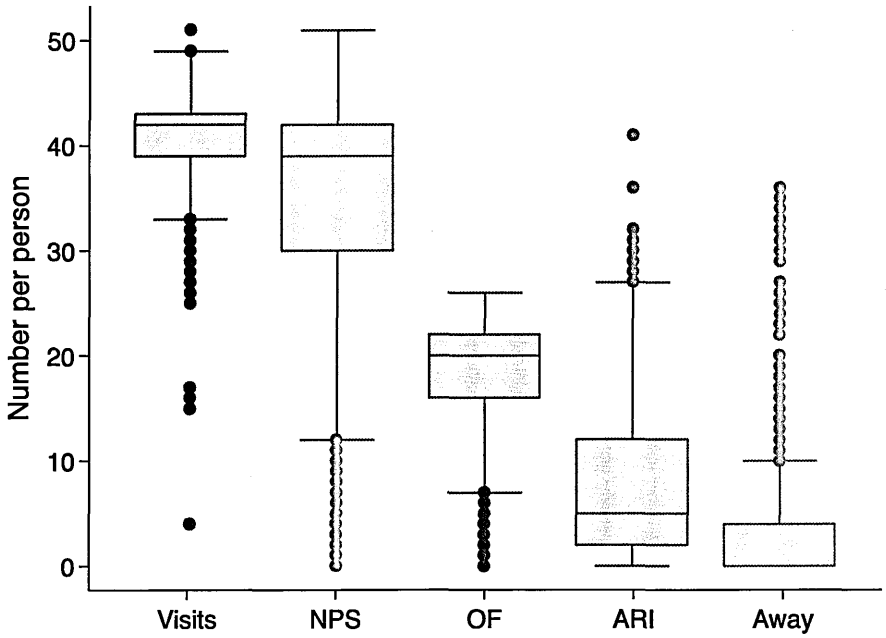


Figure 4.7: Number of home visits, NPS, OF and ARI per individual over the study period. ARI refers to number of times the individual had either blocked/runny nose, or cough or difficulty in breathing during the home or clinic visits

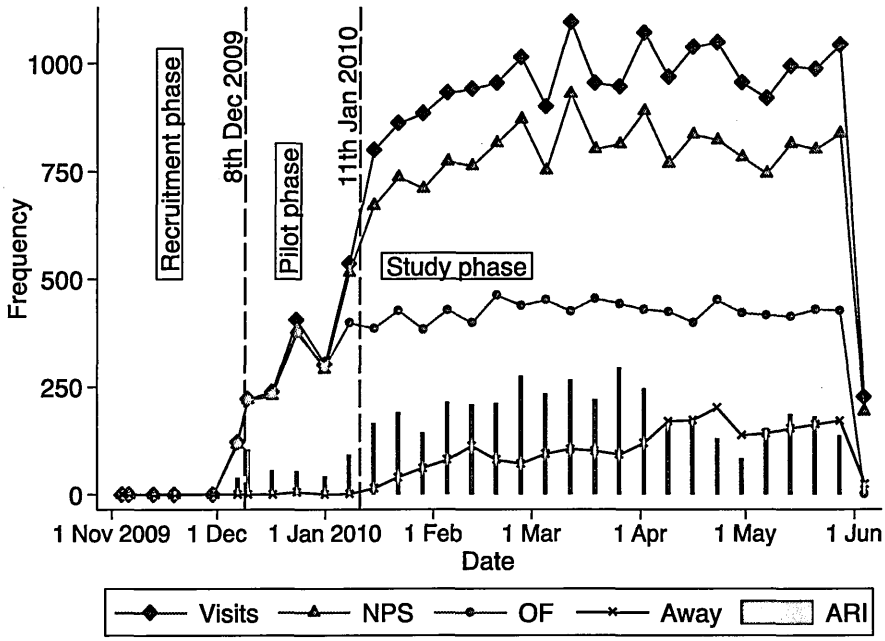


Figure 4.8: Trends in number of home visits, NPS, OF and ARI over the follow up period

4.4.9 Participant follow up and sample collections

A total of 21,566 visits were recorded during the pilot and main study phase: 21,193 (98.3%) were home visits, and the remainder clinic visits. In 2,300 (10.9%) of the home visits the participant was not available and recorded as being away, i.e. having travelled out of the village for more than 3 days. 239 (48.5%) individuals from 40 different households were reported to have been away at least once in the follow up time. The study clinic was utilized by 207 (42.0%) of all the study participants from 45 different households. A total of 16,924 NPS collections were collected and tested from 483 participants. The remaining ten (2.0%) participants (siblings (2), fathers (3), grandparent (1) and other relatives (4)) from 8 households were never sampled (NPS). The table below (Table 4.7) details the reasons for not sampling from these ten individuals. All the 493 individuals from the participant households were included in

the subsequent analyses unless otherwise stated. While the ten individuals were not tested, they could have contributed to transmission in the household and the anticipated effect of their inclusion could have minimal reduction in the estimates of attack rates since they were few.

Table 4.7: Characteristics of the participants and reasons of no NPS collections

Individual	Relations[1]	Age[2]	Reason for no NPS collection
1	Father[3]	41.6	Refused nasal sampling
2	Father	33.5	Mostly absent
3	Father	32.0	Declined nasal sampling
4	Uncle	29.4	Initially refused sampling and later travelled
5	Uncle[4]	27.5	Refused sampling citing nose bleeding
6	Uncle	21.5	Consistently left the house early and returned late from work
7	Grandfather[4]	81.7	Mostly absent
8	Sister	11.5	Away in boarding school
9	Sister[3]	8.4	Had oral cavity tumour and died in the course of the study
10	Aunt	26.5	Refused nasal sampling citing it was painful

Key: 1, relationship of the participant to study infant; 2, age in years at recruitment; 3, indicates individuals from the same household; 4, indicates individuals from the same household

4.4.10 RSV infection detections

A total of 16,924 NPS were tested by M-PCR for RSV group A and B from the 493 individuals of the 47 households. In those PCR positive, RSV group A alone was detected in 231 (1.4%) samples, group B alone in 287 (1.7%) and co-infection of

group A and B in 19 (0.1%). The two RSV groups co-circulated peaking at the same time, March to May 2010 (Figure 4.9 and Figure 4.10). A total of 205 individual episodes were detected in 179 individuals; 155 individuals had one episode, 22 had two episodes and 2 had three episodes. Household episodes were observed in 40 (85.1%) of the 47 study households (Table 4.8). Further details on RSV infections in households including spread within the households are presented in Chapter 6.

Table 4.8: Summary of RSV infections in households, individuals and NPS collections

Level	RSV A	RSV B	RSV
a) Households (47)			
No. of HH ever infected (%)	25 (53.2)	34 (72.3%)	40 (85.1)
No. of HH episodes per HH			
None	22	13	7
One	18	22	17
Two	6	8	12
Three	1	4	8
Four	-	-	2
Five	-	-	0
Six	-	-	1
b) Individuals (493)			
No. of individuals ever infected	88 (17.8)	113 (22.9)	179 (36.3)
No. of episodes			
None	405	380	314
One	81	102	155
Two	5	11	22
Three	1	0	2
No. of RSV episodes	94	124	205
Repeat infections, n (%)	7 (7.4)	11 (8.9)	26 (12.7)
c) Samples (16,924)			
RSV positive, n (%)	250 (1.5)	306 (1.8)	537(3.2)

Key: RSV, includes detection of either RSV group A and/or B; HH, household

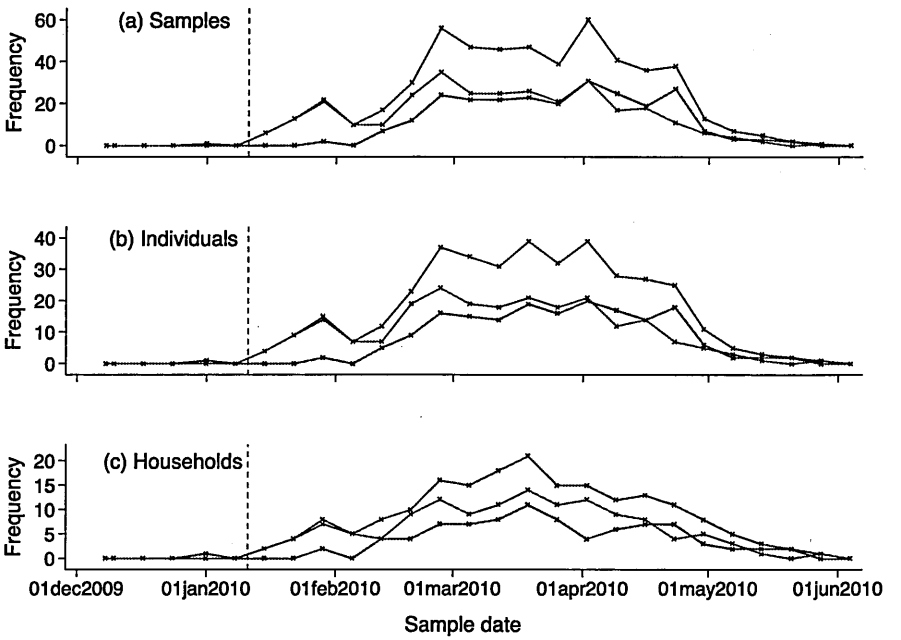


Figure 4.9: Weekly detections of RSV in NPS collections (a), individuals (b), and households (c). The navy blue, green and maroon lines denote RSV, RSV A and RSV B, respectively. The dashed light grey vertical line shows the start of the main study phase

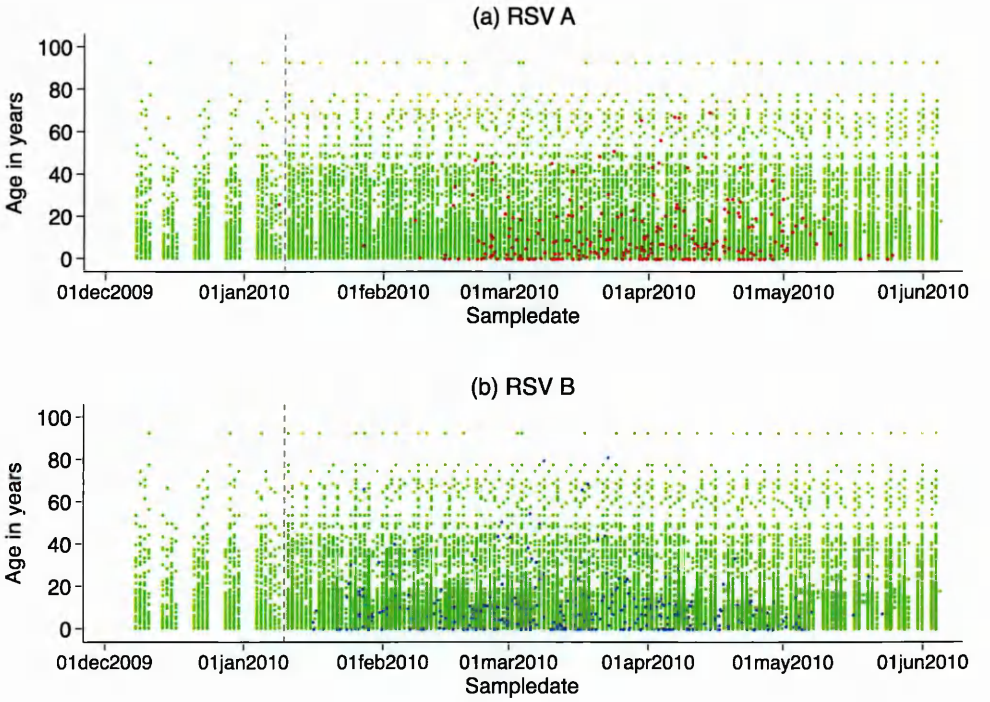


Figure 4.10: RSV A (a) and RSV B (b) detections over the study period by age in years. The green dots indicate PCR negative NPS collections while PCR positive NPS sample are red and blue dots for RSV group A and B, respectively. The dashed light grey vertical line shows the start of the main study phase

4.4.11 RSV attack rates and individual episodes

Overall the crude attack rate of the cohort was 36.3% (179/493). RSV group A infection occurred in 88 (17.8%) individuals while RSV group B was detected in 113 (22.9%) individuals. However, the risk of infection was age dependent (Figures 4.11 and 4.12 and Table 4.9). Given that the age groups mapped closely onto the categorization of the household members, the attack rates by age group and by relationship to the study infant were similar (Figure. 4.12 and 4.10). The study infants, siblings, cousins, mother, father and other householders had attack rates of 59.6%, 41.5%, 39.5%, 26.1%, 18.2% and 20.3%, respectively. The crude attack rates

among those attending school were not statistically different from those for the non-school goers (38.2% versus 35.0%, p=0.455).

In the 40 households with at least one individual diagnosed with RSV infection, the overall attack rate was 40.5% (179/442), Table 4.10. The secondary attack rates were again age-dependent with 40.9% of the household contacts aged <1year being infected. Only 11.8% of the household contacts aged 50 or more years were infected. Overall, 26.3% (114/432) of the household contacts were infected.

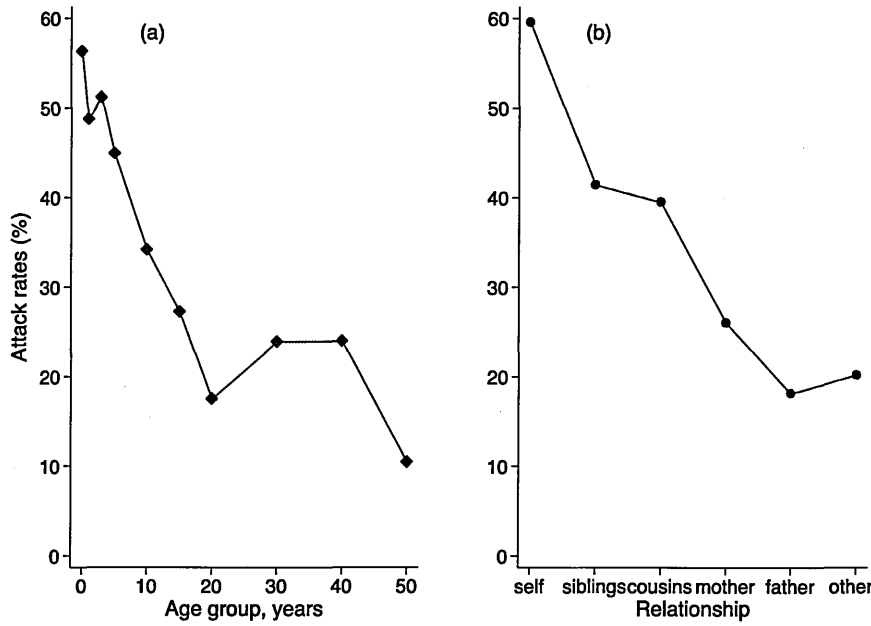


Figure 4.11: Risk of RSV infection by age at start of sampling (a) and relationship (b) to the study infant. See Table 4.9 for details on the age classes used

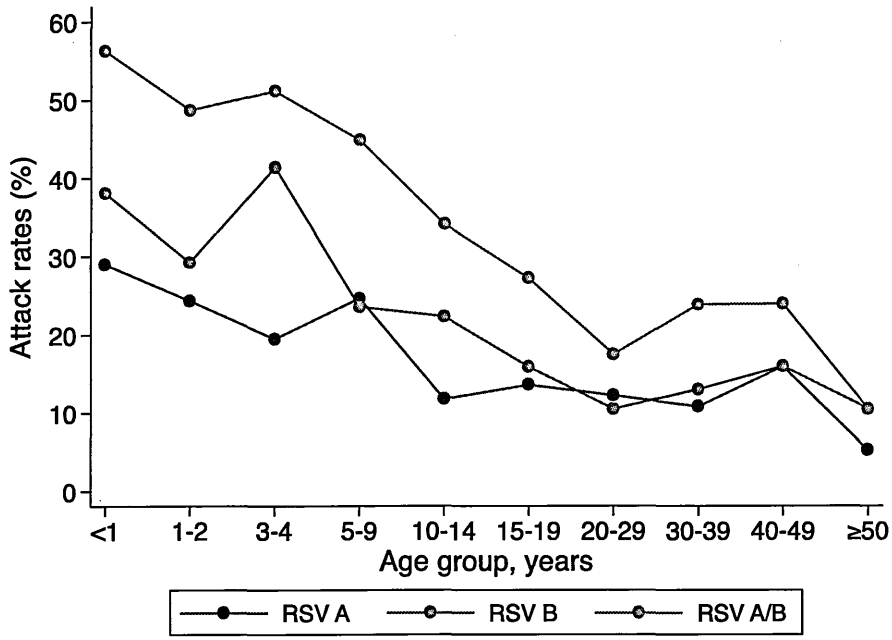


Figure 4.12: Risk of RSV infection by age at start of sampling by RSV group

Table 4.9: Crude attack rates by age, household size and relationships

			RSV A		RSV B		RSV	
Characteristics		N	n	%	n	%	n	%
Age in years at	0 - <1	55	16	29.1	21	38.2	31	56.4
start of sampling	1 – <2	41	10	24.4	12	29.3	20	48.8
	3 – <5	41	8	19.5	17	41.5	21	51.2
	5 – <10	89	22	24.7	21	23.6	40	44.9
	10 – <15	76	9	11.8	17	22.4	26	34.2
	15 – <20	44	6	13.6	7	15.9	12	27.3
	20 – <30	57	7	12.3	6	10.5	10	17.5
	30 – <40	46	5	10.9	6	13.0	11	23.9
	40 – <50	25	4	16.0	4	16.0	6	24.0
	≥50	19	1	5.3	2	10.5	2	10.5
Household size	4	37	11	29.7	18	48.6	22	59.5
	6	59	18	30.5	13	22.0	26	44.1
	8	91	2	2.2	16	17.6	18	19.8
	10	101	7	6.9	27	26.7	32	31.7
	15	97	15	15.5	13	13.4	27	27.8
	20	108	35	32.4	26	24.1	54	50.0
Relationship to	Self	47	13	27.7	21	44.7	28	59.6
the study infant	Sibling	164	34	20.7	41	25.0	68	41.5
	Cousin	124	21	16.9	31	25.0	49	39.5
	Mother	46	8	17.4	7	15.2	12	26.1
	Father	33	3	9.1	4	12.1	6	18.2
	Other	79	9	11.4	9	11.4	16	20.3

Table 4.10: Crude and secondary attack rate of RSV in households according to age at start of sampling

Age groups	All households			PCR positive households			Secondary attack rate		
	No. in study	*No. Infected*	*%*	*No. in study*	*n*	*%*	*Contacts*	*n*	*%*
<1	55	31	56.4	48	31	64.6	44	18	40.9
1-2	41	20	48.8	35	20	57.1	35	17	48.6
3-4	41	21	51.2	36	21	58.3	35	16	45.7
5-9	89	40	44.9	84	40	47.6	83	24	28.9
10-14	76	26	34.2	69	26	37.7	69	12	17.4
15-19	44	12	27.3	41	12	29.3	41	7	17.1
20-29	57	10	17.5	51	10	19.6	48	8	16.7
30-39	46	11	23.9	39	11	28.2	39	6	15.4
40-49	25	6	24.0	22	6	27.3	21	4	19.0
≥50	19	2	10.5	17	2	11.8	17	2	11.8
Overall	493	179	36.3	442	179	40.5 %	432	114	26.3

Key: PCR, polymerase chain reaction

4.4.12 Subclinical RSV episodes

Of the 205 RSV individual episodes, 87 (42.4%) were not associated with clinical symptoms of acute respiratory illness such as cough, runny nose or blocked nose or difficulty in breathing. The pattern was not different by RSV groups, RSV group A (35/81, 43.2%) and group B (50/110, 45.5%). Coinfections of RSV A and B were linked with fewer asymptomatic infections (2/14, 14.3%) (Fisher's exact P = 0.046 comparing to the mono-group infections). The proportion of sub-clinical RSV infections increased with age (see Table 4.11). Only 3/37 (8.1%) of RSV episodes among infants (<1 year) were asymptomatic and that increased across the ages to 36/45 (80.0%) among the adults (≥15 years). RSV episodes associated with male gender, presence of other viruses, and individuals not attending school were more likely to be asymptomatic compared to their counterparts. The difference on risk of subclinical infections by gender was statistically significant only in participants older than 12 years.

Table 4.11: Characteristic of subclinical RSV episodes

			Asymptomatic infections		
Characteristics	**Categories**	**N**	*n*	*%*	*P value*
Age in years at start	0 – <1	37	3	8.1	<0.0001
of sampling	1 – <5	49	10	20.4	
	5 – <15	74	38	51.4	
	15 – <40	36	29	80.6	
	≥40	9	7	77.8	
Relationships of the	Self	34	1	2.9	<0.0001
households members	Siblings	79	31	39.2	
to the study infant	Cousins	55	24	43.6	
	Mothers	14	11	78.6	
	Fathers	6	4	66.7	
	Others	17	16	94.1	
Gender	Female	116	59	50.9	
	Male	89	28	31.5	0.005
Presence other	No	122	63	51.6	
viruses[2]	Yes	83	24	28.9	0.001
Other co-infection	None	122	63	51.6	Ref.
viruses	Adenovirus	23	10	43.5	0.23
	Coronaviruses	19	6	31.6	0.092
	Rhinoviruses	19	3	15.8	0.001
	Mixed	22	5	22.7	0.001
During an RSV	No	42	26	61.9	
outbreak in HH	Yes	163	61	37.4	0.004

Participant in school	No	137	47	34.3	
	Yes	68	40	58.8	0.001
Repeat infections	First	179	73	40.8	
	Second[3]	26	14	53.9	0.208
Infecting RSV group	A	81	35	43.2	Ref.
	B	110	50	45.5	0.771
	A & B	14	2	14.3	0.072

Key: 1, nonparametric test for trend across the age groups; 2, Presence of any of the other viruses (rhinoviruses, adenoviruses, and coronaviruses) during the RSV episode; 3, includes the second and the third RSV repeat infections per individual; N, number of RSV infection episodes; Ref, reference category

4.4.13 The Ct values distribution for the viral targets in the real time M-PCR assay

A total of 16,924 nasal samples from both the pilot and study phase of the household study were screened. In 1226 (7.2%) of the NPS screened, a Ct value for either RSV group A or B was recorded. Of the common respiratory pathogens screened for, a bimodal distribution of the Ct values was observed in RSV group A and B, NL63 and 2293 (Figure 4.13). Ct values for adenovirus, OC43 and rhinovirus had a unimodal distribution. The mean Ct values increased with age (Table 4.12). Individuals aged 5 years or more had NPS collections with higher mean Ct values than the younger counterparts. This phenomenon was observed for adenoviruses, coronaviruses and rhinoviruses (see Figures and Tables in Appendix R).

Table 4.12: Comparison of Ct values for RSV group A and B by age

	RSV group A				RSV group B			
Age[1], years	*n*	*Mean*	*95% CI*	*P-value*	*n*	*Mean*	*95% CI*	*P-value*
<1 y	125	32.3	31.5 – 33.2	Ref	117	28.3	27.2 – 29.4	Ref.
1-<5 y	164	33.5	31.4 – 35.6	0.056	98	29.1	26.4 – 31.9	0.332
5 -<15 y	260	34.6	32.6 – 36.6	<0.001	161	31.8	29.2 – 34.3	<0.001
15-<40 y	182	35.8	33.7 – 37.8	<0.001	66	32.9	29.9 – 35.9	<0.001
≥40 y	74	35.4	33.0 –37.7	<0.001	19	32.7	28.6 – 36.8	0.004

Key: 1, age groups in years (y); n, the number of samples with a Ct value; CI, confidence interval; Ref, reference group

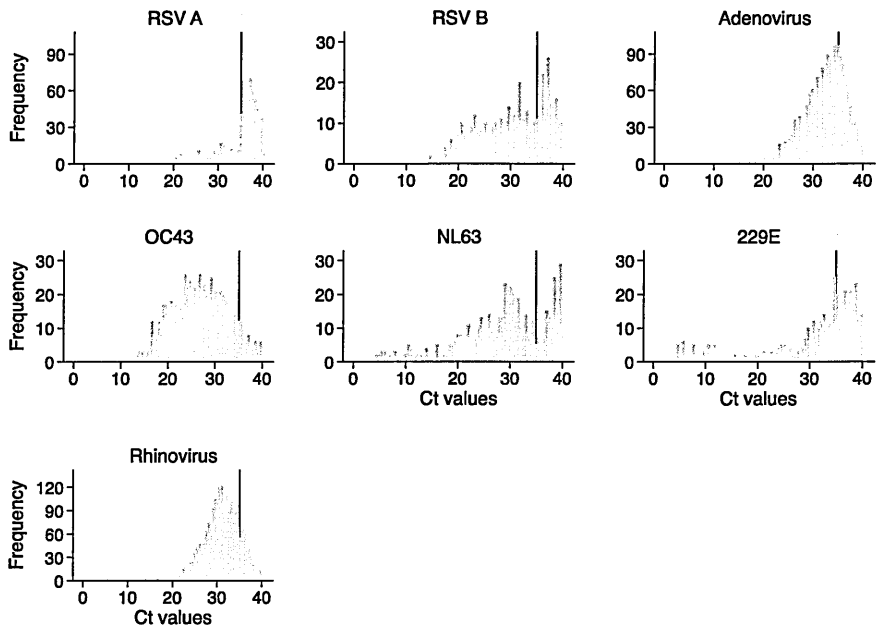


Figure 4.13: Ct values distribution for the most prevalent respiratory viruses detected. OC43, NL63 and 229E are strains of human coronaviruses

4.4.14 RSV 'suspected' repeat infections

Twenty-four (13.4%) participants experienced more than two RSV infection episodes. Most (17/24, 70.8%) of the repeat infections were with homologous RSV group (group A (6) and group B (11)) while 29.2 % (7/24) were with heterologous group (Table 4.13). The median (IQR) interval between homologous 'reinfections' was shorter (24.5 (21 - 37) days) than for heterologous 'reinfections' 54 (47 - 62) days), Wilcoxon rank sum test, p=0.0006. Data on genetic diversity of the infecting RSV variants based on sequencing of the long ectodomain of the RSV attachment gene are presented in Chapter 6.

Table 4.13: Identity of the infecting RSV group of first and second episodes

		Second episodes[1]					
	RSV groups	A			B		
		Interval[2]	n	%	Interval[2]	n	%
First episodes	A (N=7)	25 (20 – 31)	6	85.7	49 (-)	1	14.3
	B (N=17)	59 (45 – 62)	6	35.3	30.1 (21 – 38)	11	64.7

Key: 1, two individuals had three episodes but only the second is included in the tabulation. One individual had the three RSV group A episodes while the other had RSV group B for the first and second and RSV group A for the third episodes; 2, median (IQR) interval between episodes in days

4.4.15 The most prevalent respiratory pathogens during the study period

In line with the strategy to establish the most prevalent respiratory viruses in the community during the study period, NPS collections from six households were screened for the full range of respiratory viruses targeted by the M-PCR assay. The baseline characteristics of these households are presented in Table 4.14 and Figure 4.14. Of the 2,644 samples screened, 279 (10.6%), 250 (9.5%), 214 (8.1%), 151 (5.7%), 61 (2.3%), 19 (0.7%), 10 (.4%), and 9 (0.3%) were positive for rhinoviruses, adenoviruses, coronaviruses (OC43, NL63, or 229E), RSV (group A or B), PIV (type 1, 2, 3 or 4), *Mycoplasma pneumoniae*, influenza viruses (type A, B or C) and hMPV, respectively (Figure 4.15). In addition to RSV group A and B, rhinoviruses, adenoviruses, and coronaviruses (OC43, NL63 and 229E) were the most prevalent respiratory viruses and taken forward for screening from the rest of NPS collections.

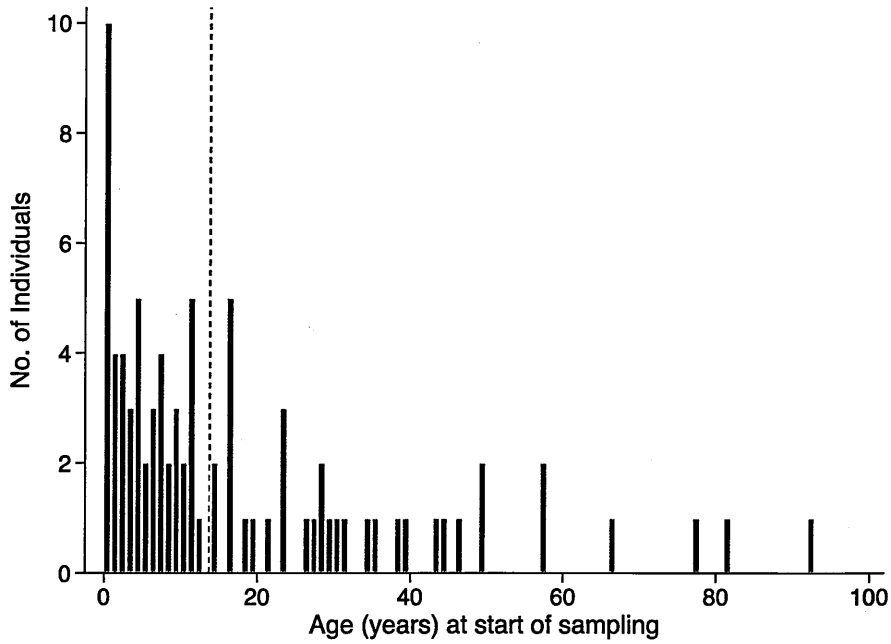


Figure 4.14: Age distribution of individuals in the six households screened for a full range of respiratory viruses. Red dotted line shows the mean age in years

Table 4.14: Baseline characteristics of 83 individuals from six households tested for all virus targets

Characteristic	**Statistic**
Household size, Median (IQR)	10.5 (5 – 15)
No. (%) of school-going children	21 (25.3)
Number (%) male individuals	31 (37.4%)
Median (IQR) age, at start of sampling	10.3(3.5 – 27.5)
Age groups, in years[1]	
<1y	10 (12.1)
1– 4y	16 (19.3)
5– 14y	24 (28.9)
15– 39y	22 (26.5)
>=40y	11 (13.3)

Key: IQR, interquartile range; 1, age at start of sampling

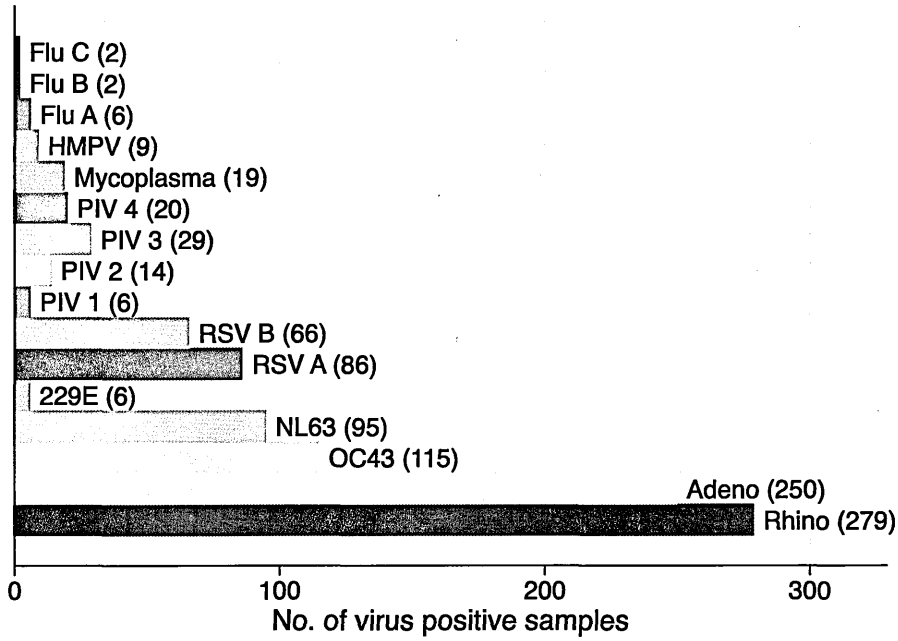


Figure 4.15: Frequency distribution of the detected respiratory viruses in 2644 NPS collections from 83 individuals of the six households with full respiratory screen. Rhino, rhinoviruses; Adeno, adenoviruses; OC43, NL63 and 229E are strains of coronaviruses; RSV A and B, respiratory syncytial virus group A and B; PIV 1, 2, 3 and 4, parainfluenza type 1, 2, 3, and 4; Mycoplasma; mycoplasma pneumoniae; HMPV, human metapneumoviruses; Flu A, B and C, influenza type A, B and C

4.4.16 Common respiratory viruses in the study cohort

The 16,924 NPS collections were tested for rhinovirus, adenovirus, coronaviruses (OC43, NL63 and 229E) as well as RSV (group A and B). Of these samples, 4326 (25.6%) had one or more viruses detected. Rhinovirus and coronaviruses were detected in at least one person from the 47 households while adenovirus and RSV were detected in 45 (95.7%) and 40 (85.1%) households, respectively, Table 4.15 and Figure 4.16. The individual attack rates were 93.4%, 80.1%, 71.6%, 61.5% and 37.1% for any virus, rhinoviruses, coronaviruses, adenovirus and RSV, respectively. These

viruses circulated throughout the study period as shown in Figure 4.17. On average, each individual had three unique viral infections over the ~6-month study period, Figure 4.18.

Table 4.15: Virus detections in households, participants and NPS collections from the 483 participants from the 47 study households

Virus	Detected respiratory viruses in:					
	Households (N=47)		Individuals (N=483)[1]		NPS (N=16924)	
	n	*%*	*n*	*%*	*n*	*%*
Any virus	47	100.0	451	93.4	4326	25.6
Rhinovirus	47	100.0	387	80.1	1760	10.4
Coronaviruses	47	100.0	346	71.6	1268	7.5
OC43	*44*	*93.6*	*212*	*43.9*	*645*	*3.8*
NL63	*33*	*70.2*	*163*	*33.7*	*418*	*2.5*
E229	*30*	*63.8*	*119*	*24.6*	*241*	*1.4*
Adenovirus	45	95.7	297	61.5	1232	7.3
RSV	40	85.1	179	37.1	537	3.2
Group A	*25*	*53.2*	*88*	*18.2*	*250*	*1.5*
Group B	*34*	*72.3*	*113*	*23.4*	*306*	*1.8*

Key: 1, Excludes 10 participants who were never sampled.

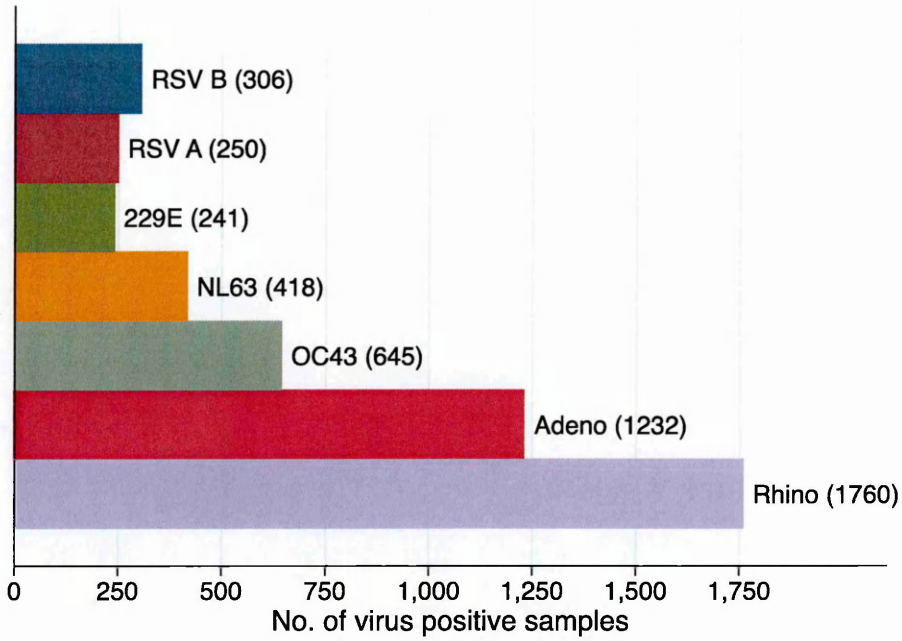


Figure 4.16: Frequency distribution of the common respiratory viruses from the 47 study households. Rhino, rhinoviruses; Adeno, adenoviruses; OC43, NL63 and 229E are strains of coronaviruses; RSV A and B, respiratory syncytial virus group A and B

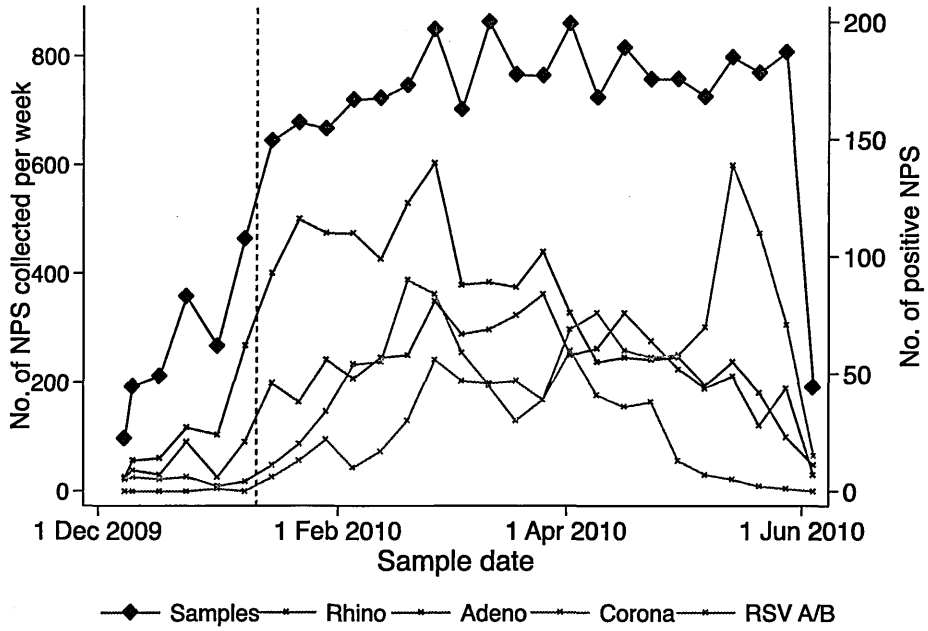


Figure 4.17: Weekly detections of the most prevalent viruses in NPS collections from the 47 study households. The vertical dashed line denotes the start of the main study period. Rhino, rhinoviruses; Adeno, adenoviruses; corona, coronaviruses; RSV A/B, respiratory syncytial virus group A and/or B

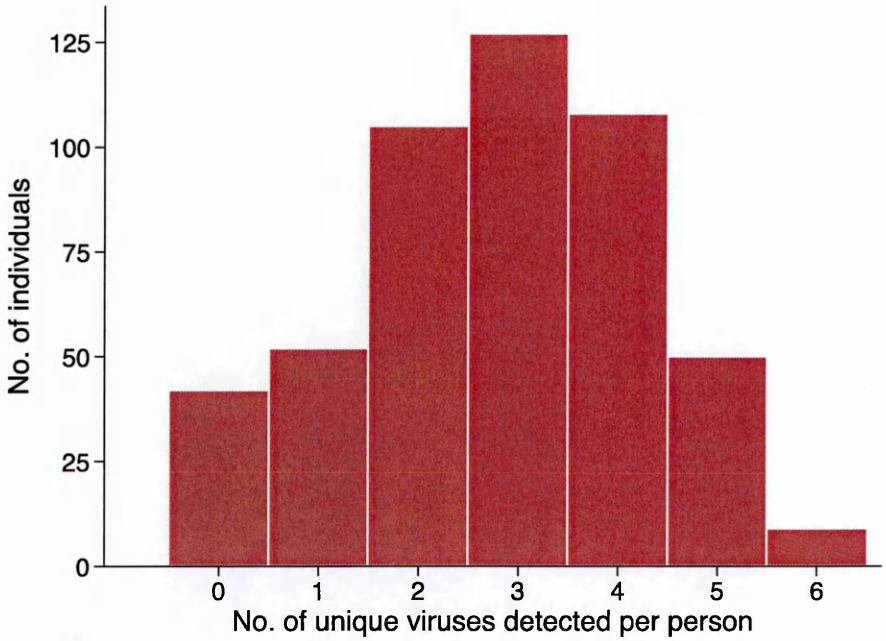


Figure 4.18: Number of different viruses detected per person over the study period

4.4.17 Prevalence of viral co-infections with RSV

About a quarter (133) of the RSV positive NPS collections were also positive for other respiratory viruses: adenoviruses (10.1%), rhinoviruses (9.5%) and coronaviruses (8.6%). Similar frequencies of co-infections were noted for the RSV infection episodes (see Table 4.16).

Table 4.16: Respiratory viruses co-detected with RSV in NPS collections

a) Co-detection of viruses in same NPS collection	***RSV A***		***RSV B***		***RSV***[1]	
	n	*%*	*n*	*%*	*n*	*%*
Any virus	68	27.2	72	23.5	133	24.8
Adenoviruses	31	12.4	23	7.5	54	10.1
Rhinoviruses	17	6.8	38	12.4	51	9.5
Coronaviruses	32	12.8	18	5.9	46	8.6
OC43	23	9.2	8	2.6	31	5.8
NL63	5	2.0	7	2.3	9	1.7
229E	4	1.6	3	1.0	6	1.1
b) Viral co-infections during an RSV infection episodes[2]						
No. of co-infection[2]	53	57.0	73	61.3	122	59.5
Any co-infection	40	43.0	46	38.7	83	40.5
Adenovirus	9	9.7	14	11.8	23	11.2
Coronaviruses	11	11.8	11	9.2	19	9.3
Rhinovirus	7	7.5	12	10.1	19	9.3
Mixed[3]	13	14.0	9	7.6	22	10.7

Key: 1, A total of 537 samples were positive for RSV (250 with RSV group A and 306 with group B) 2, total of 205 RSV episodes (94 with group A and 124 with group B); 2, co-infection defined as detection of another virus during the period of an RSV episode but not necessarily in the same sample; 3, two or more viruses detected during the RSV episode

4.4.18 Prevalence of viruses stratified by clinical status

The selection of the samples for this analysis was staggered. First, 229 samples collected when 41 infants were symptomatic for ARI were selected. This samples covered only the period between 1st January and 31st March 2010. This first batch

was also aimed at identifying the common respiratory viruses in the cohort hence the focus in the susceptible the age group i.e. under one year of age. The second batch included 343 samples from 163 participants with ARI at time of collection selected at random from the full study period. For comparison purposes, a further 446 samples collected during asymptomatic times from 283 participants randomly selected. A full respiratory screen was done in all the 1018 NPS collections to identify the prevalence of viral detections by ARI status: 572 samples from individuals who had symptoms of ARI and 446 samples from individuals without ARI at the time of sample collection. The age distributions of the 362 participants involved in this analysis are shown in Table 4.17. The two groups (ARI and non-ARI) had statistically significant differences by age and sex composition. This is largely to the poor sample selection criteria adopted.

Table 4.17: Baseline characteristics of the 362 participants

Characteristic		**ARI (N=201)**	**No ARI (N=283)**	**P-value**
Age in years, mean (95% CI)		11.4 (9.4 – 13.4)	16.9 (14.9 – 18.9)	0.0002
Age at the time of	<1 y	41 (20.4)	34 (12.0)	0.001
sample collection, n	1-<5y	46 (22.9)	49 (17.3)	
(%)	5-<10y	39 (19.4)	54 (19.1)	
	10-<20y	38 (18.9)	58 (20.5)	
	20-<40y	26 (12.9)	56 (19.8)	
	≥40	11 (5.5)	32 (11.3)	
Male gender, n (%)		96 (47.8)	102 (36.0)	0.0098

Key: ARI, acute respiratory illness

Table 4.18: Frequency distribution of respiratory virus detection by ARI status

	ARI (N=572)		No ARI (N=446)		
Detected viruses	*n*	%	*n*	%	**P value[1]**
Rhinoviruses	158	27.6	16	3.6	<0.0001
Adenoviruses	70	12.2	7	1.6	<0.0001
RSV	71	12.4	17	3.8	<0.0001
Group A	38	6.6	9	2.0	0.0005
Group B	33	5.8	8	1.8	0.0014
Coronaviruses	50	8.7	20	4.5	0.0077
OC43	38	6.6	14	3.1	0.0117
NL63	6	1	5	1.1	1
229E	6	1	1	0.2	0.144
Parainfluenza viruses	22	3.8	0	0	-
Type 1	0	0	0	0	-
Type 2	8	1.4	0	0	-
Type 3	12	2.1	0	0	-
Type 4	2	0.3	0	0	-
hMPV	3	0.5	0	0	-
Influenza viruses	1	0.2	3	0.7	0.325
Type A	0	0	0	0	-
Type B	0	0	1	0.2	-
Type C	1	0.2	2	0.5	0.585
Mycoplasma pneumoniae	0	0	0	0	-

Key: 1, from chi- square test except for NL63, 229E and Influenza viruses; ARI, acute respiratory illness; Rhino, rhinoviruses; Adeno, adenoviruses; OC43, NL63 and 229E are strains of coronaviruses; RSV A and B, respiratory syncytial virus group A and B

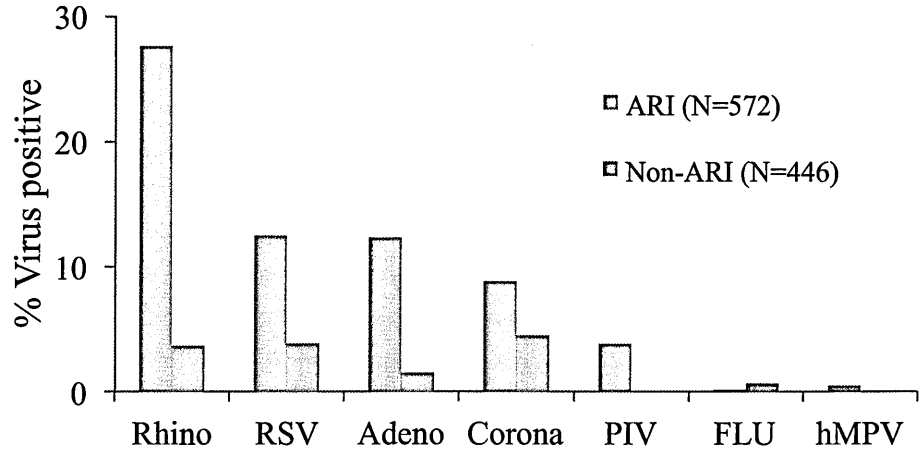


Figure 4.19: Prevalence of viruses by ARI status. Rhino, rhinoviruses; Adeno, adenoviruses; corona, coronavirures (OC43, NL63 and 229E); RSV, respiratory syncytial viruses (group A and B); PIV, parainfluenza viruses (type 1, 2, 3 & 4); FLU, Influenza viruses (type A, B and C); hMPV, human metapneumoviruses

4.4.19 Oral Fluid sensitivity in detecting RSV by the M-PCR assay

In assessing the diagnostic performance of oral fluid, 59 randomly selected OF samples with paired NPS that were RSV positive were screened for RSV by M-PCR. Overall, 23 OF samples were reported positive giving a sensitivity of 39. 0 (95% CI; 26.5 – 52.6), Table 4.19. The sensitivity was similar for detection of the two RSV groups (36.1% and 42.9% for group A and group B, respectively, p=0.583). The sensitivity of OF samples increased where the paired NPS had low Ct values (56.3% for Ct values ≤25), were from young ages (61.5% for under 5 years) or were from symptomatic episodes (52.2% when associated with ARI). On average the Ct values from the OF samples were lower than those for NPS (28.14 vs. 36.14 respectively; Paired t-test P value <0.001.) The pattern did not differ on limiting the analyses to samples positive by both sampling method, Table 4.20. Use of an alternative

screening test (Fast-track Diagnostics, Luxembourg) did not improve the diagnostic performance of OF in detection of RSV (Data shown in Appendix S).

Table 4.19: Sensitivity of oral fluid in detection of RSV by M-PCR

			No. of RSV	Sensitivity		
Characteristic	**Categories**	**N**	**Positives**	*%*	*95% CI*	**P-value[1]**
Overall	All	59	23	39.0	26.5 – 52.6	-
NPS Cts	>30 – 35	23	7	30.4	13.2 – 52.9	0.107
	>25 – 30	25	10	40.0	21.1 – 61.3	0.309
	≤25	16	9	56.3	29.9 – 80.2	Ref.
RSV group	RSV A	36	13	36.1	20.8 – 53.8	0.583
	RSV B	28	12	42.9	24.4 – 62.8	
Age (years)	<5 y	26	16	61.5	40.6 – 79.8	0.002
	≥5 y	33	7	21.2	9.0 – 38.9	
Age (years)	<1 y	9	5	55.6	21.2 – 86.3	Ref.
	1 – <5 y	17	11	64.7	38.3 – 85.8	0.692
	5 – <10 y	13	4	30.8	9.1 – 61.4	0.3842
	≥10 y	20	3	15.0	3.2 – 37.9	0.067
Gender	Female	29	12	41.3	23.5 – 61.1	0.711
	Male	30	11	36.7	20.0 – 56.1	
ARI	Yes	23	12	52.2	30.6 – 73.2	0.097
	No	36	11	30.6	16.3 – 48.1	

Key: ARI, acute respiratory illness; 1, chi-square test except for characteristic with 5 or less RSV positive samples; Ref; reference category

Table 4.20: Comparison of NPS and OF Ct values for detection of RSV by M-PCR

	NPS		OF		
Detection	*Mean Ct*	*95% CI*	*Mean Ct*	*95% CI*	*P value*
a) Both NPS and OF positive (N=25)					
RSV	26.9	24.8 – 29.0	30.9	29.3 – 32.5	0.001
Group A (16)	18.9	12.6 – 25.1	31.1	28.8 – 33.4	0.009
Group B (9)	19.2	9.9 – 28.5	30.5	28.4 – 32.5	0.023
b) All the screened sample (N=59)					
RSV	28.1	26.9 – 29.4	36.1	34.8 – 37.5	<0.001
Group A (38)	28.6	27.0 – 30.3	37.2	35.6 – 38.8	<0.001
Group B (28)	26.3	24.6 – 28.0	37.7	36.0 – 39.4	<001

Key: Comparison in (a) included only samples that were PCR positive (Ct value ≤35) for both NPS and OF while (b) included all samples (recoding the Ct values for OF to 40 if the Ct was indeterminate); N, number of samples tested

4.4.20 RSV-specific antibody profiles in OF samples

Samples of OF from a household with five members consisting of the study infant (1 month old at start of sampling), a sister (2 years), two brothers (6 and 9 years) and mother (29 years) were tested for RSV-specific antibody titers using ELISA method as described in Chapter 3. All except the study infant and the mother had low antibody titres at the start of the study (Figure 4.20). There was measurable antibody rise for the four individuals with PCR-detected RSV infections. However, additional peaks were observed which were not associated with RSV detections by M-PCR.

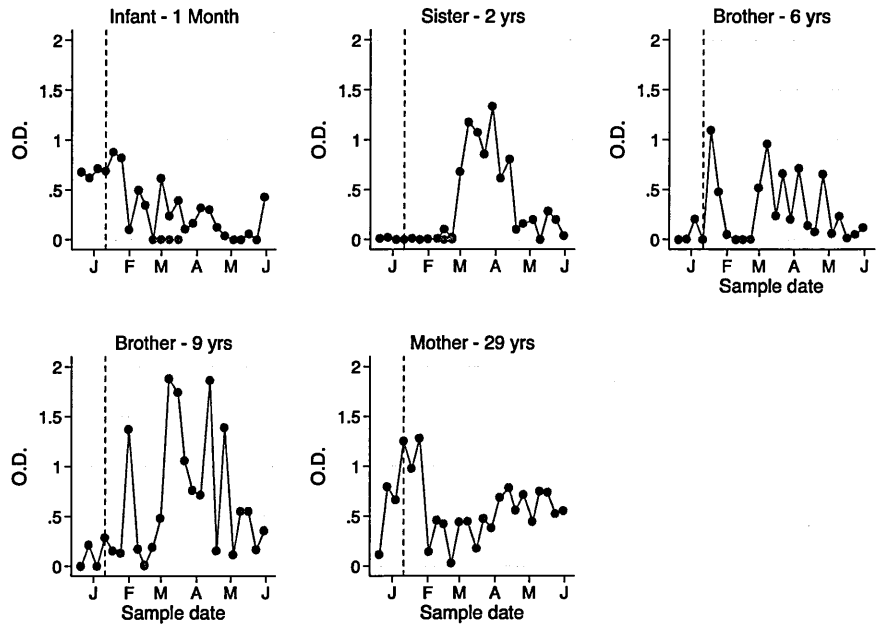


Figure 4.20: Antibody profiles for five members of the same household with consistent weekly OF collections. The navy blue line shows the RSV-specific antibody profiles while the maroon filled circles show the PCR detections of RSV in NPS collections. The vertical dashed line denotes the start of the main study phase. O.D, adjusted optical density; age shown was at the start of sampling

4.5 Discussion

We report infection data from 47 households with 493 individuals followed for 24 weeks during a single RSV epidemic. A total of 16924 deep nasopharyngeal swabs were collected twice-a-week regardless of symptoms and screened for a range of respiratory viruses including RSV. The detailed infection data provided adequate data to allow investigation of transmission of the viruses within the household. Data on transmission patterns within the household are presented in Chapter 6.

The study was carefully implemented and monitored (as detailed in Chapter 3) yielding a high retention rate of 78% of the recruited households and high compliance in sample collection of over 80%. About 16,500 (about 84% of the expected) NPS were collected over the 24-week study period. Compliance in sample collection was higher among the study infants, older children and their mothers (over 80%) in comparison to their fathers and other adults in the households (~ 65%). Previous family studies have reported poor compliance from male parents and our compliance of about 65% is relatively high (Hall *et al.* 1976; Okiro 2007). It is also likely that the provision of free health care at the study clinic and monthly home visits by the clinician for health check contributed to the high acceptability and compliance in sampling.

RSV seasonality as captured from the inpatient RSV surveillance at KDH paediatric wards, guided the start and end of the household study. KDH is the main inpatient facility in the district and it has been estimated that over 60% of the paediatric admission arise from the area under demographic surveillance which covers Matsangoni location, the site of the current study (Moisi *et al.* 2011; Scott *et al.* 2012). Based on the inpatient data, the household study was initiated in time before the RSV epidemic began covering the 24 weeks of RSV circulation in the community. There was a lag of about 6 weeks in detection of RSV cases in Matsangoni relative to the District hospital auguring well for the present study. This could be attributed to variability in temporal and/or spatial patterns of RSV infections throughout the KDH catchment area. It would be worthwhile to assess whether RSV spreads systematically within the District or by pure chance causing the local epidemics. For study purposes, the lag allowed the pilot phase to be completed in time with the study phase

conveniently capturing most of the local RSV epidemic. It is no surprise the RSV epidemic in Matsangoni was shorter relative to KDH for it is a function of susceptible population size within the study location – the household study captured cases within a small area of Matsangoni location only.

Co-circulation of both the RSV Group A and B during the 2009/2010 RSV season was captured in the two concurrent studies (Matsangoni household and KDH inpatient study). This was useful in distinguishing infection episodes and transmission events within households as reported here and in Chapters 5 and 6.

The study cohort had appreciable RSV attack rates with (85.1%) of the household and (36.3%) of the study participnats being infected. Existing data suggests infants are most susceptible to RSV infection (Glezen *et al.* 1986; Nokes *et al.* 2009) and in the current study; they had high attack rates of about 60%. The risk of infection declined with age with a slight increase in attack rates for 30-49 year olds which could be related to higher contact rates between young children and parents particularly mothers. Mothers had higher attack rate relative to the fathers (26.1% vs. 18.2). Similar pattern of attack rates by age were reported in the Rochester family study (Hall *et al.* 1976) and could be attributed to acquisition of partial protective immunity following virus exposure in successive RSV epidemics (Glezen *et al.* 1986). Our estimates of attack rates were higher relative to the Hall et al study, which could be explained by methodological differences. In the current study, deep nasopharyngeal samples were collected over a longer period (24 vs. 8 weeks) and screened with a more sensitive molecular method (relative to culture in the Hall study). Other family studies such as the Seattle virus-watch studies (Cooney *et al.* 1972; Cooney *et al.* 1975) and Berglund serologic study (Berglund 1967) of family members of children

admitted with RSV infection reported considerable attack rates in older ages (reviewed in Chapter 2). The role of the older children and adults in transmission is further investigated in Chapters 5 and 8. Smaller households with 4 – 7 members had high attack rates relative to large households probably due to close contact network in these households: all were nuclear family units compared to households with more than 8 members which comprised of several nuclear family units (45.5% had two or more families).

About half (42.4%) of the reported RSV episodes were asymptomatic occurring mainly in older children and adults. These findings have implications for the interpretation of results from studies relying on symptoms for RSV surveillance. Such studies would underestimate the RSV infection rates particularly in older ages groups and are not suitable for identifying links of infection spread. Clinical severity during the first and second RSV episode was similar. Males were more likely to be asymptomatic relative to females indicating a possible reporting bias.

Rhinoviruses, adenovirus and coronaviruses were the most prevalent respiratory viruses co-circulating with RSV. It is surprising very few cases were positive for influenza viruses despite the scare of pandemic influenza during the course of the study. Previous studies have reported co-circulation of hMPV with RSV, which was not obvious in this study (Semple *et al.* 2005). Over 90% of the study participants had at least one virus detected and on average three different viral infections over the six-month period of the study. Co-infections were not uncommon with about 40% of the RSV episodes being associated with coronaviruses, rhinoviruses, or adenoviruses. A long-term (at least one year) community surveillance study would offer a more

comprehensive data on seasonality of these respiratory viruses and their temporal association with RSV.

More viruses were detected in samples collected during a concurrent acute respiratory illness relative to when no ARI symptoms were present. This pointed to an aetiologic link of the viruses, however, a properly designed nested case-control study with clearer definition of ARI and proper matching of randomly selected controls would provide better data.

Poor diagnostic performance of OF in detecting of RSV by M-PCR and ELISA limited the use of OF to supplement the RSV infection data. Few studies have evaluated the sensitivity of OF but available data suggest higher sensitivity of about 73% relative to nasopharyngeal aspirate (von Linstow *et al.* 2006). Our estimate of 39% was generally low. However, the sensitivity was better in samples with lower Ct value (75% when Ct value was <25 in the paired NPS) or when symptomatic (52%) and in young age groups (61.5% in <5years old). The above observations suggest OF would be useful in settings associated with high viral load like hospital inpatient settings or studies targeting children in their first few years of life. Further work is recommended to explore whether use of alternative molecular screening methods such as RNA UltraSense (One-Step Quantitative RT-PCR System, optimised for low viral load detections, Applied Biosystems) to improve the sensitivity. Collection of OF was acceptable to most of the study participants resulting in high compliance of over 90% as shown in Chapter 3. Even though only OF samples from one household were screened, the observed RSV-specific antibody profiles were qualitatively less discriminatory in detecting additional RSV infections. However, screening of more OF samples is justified before a conclusive assessment is made.

CHAPTER FIVE

5 Who Brings RSV Infection into Households and Who Infects the Infants

5.1 Introduction

Human respiratory syncytial virus (RSV) is a major cause of childhood acute lower respiratory tract infection (LRTI) worldwide (Nair *et al.* 2010). The virus is characterised by seasonal outbreaks usually, though not always, occurring annually, (Waris 1991; White *et al.* 2005; White *et al.* 2007), and is transmitted efficiently in early life: approximately 60% of newborns are infected during their first epidemic (Glezen *et al.* 1986). Infection can occur in young infants despite the transfer of maternally derived passive antibodies (Ochola *et al.* 2009). The risk of severe respiratory disease following RSV infection is highest in infants, particularly those under 6 months of age (Ohuma *et al.* 2012). Together these factors lead to the observation of the vast majority of severe RSV disease cases, including seasonal congestion in hospital admissions, comprises infants and especially young infants (Vardas *et al.* 1999; Hussey *et al.* 2000; Nokes *et al.* 2009).

Set in this context, vaccine development, in particular live attenuated virus vaccines, has primarily targeted infants aged under 3 months. However, no human RSV vaccine has yet been licensed. The early infant presents severe obstacles to vaccinologists in the form of immunological immaturity, presence of maternal RSV-specific antibodies, and balancing between immunogenicity and the risk of upper respiratory tract congestion associated with live vaccines (Karron *et al.* 2005). Alternative strategies for RSV vaccination have therefore been proposed (Anderson *et al.* 2013), including delaying delivery to an older age (Nokes and Cane 2008), for

which there is empirical support. Live attenuated vaccines have been found safe and immunogenic in seronegative children 6 months of age and over (Gonzalez *et al.* 2000; Karron *et al.* 2005; Wright *et al.* 2007); sub-unit RSV vaccines boost protective antibodies in previously infected individuals (Paradiso *et al.* 1994; Englund *et al.* 1998; Power *et al.* 2001; Langley *et al.* 2009) and 40%-60% of RSV associated community severe and hospitalised disease occurs in children 6 months and above (Nokes *et al.* 2008; Hall *et al.* 2009; Nokes *et al.* 2009). Delaying vaccination of young infants to 6 months of age would fail to directly protect those most at risk of severe RSV disease. However, the risk of severe disease arising from RSV infection falls rapidly with increasing age beyond 6 months (Ohuma *et al.* 2012), probably as a result of physiologically maturation leading to increased airway size (Machata *et al.* 2010). Thus significant reduction of infection in young infants during their first RSV season would have a big impact on RSV associated mortality and morbidity. Delayed first infection of infants could move them to an older age group where the risk of RSV is significantly lowered. Consequently, a natural question is whether vaccination of older (>6 months) individuals in a community can provide additional, indirect benefit, by preventing them infecting the vulnerable early infant. This could be achieved by either targeted interruption of chains of transmission leading to infant infection, or reducing virus circulation within the community that eventually leads to a more indirect lowering of the risk of infant exposure (Fox *et al.* 1971). The potential impact and design of vaccination programmes aimed at delaying infant infection will be intimately linked to understanding the source of the infection for infants, which is the subject of the current chapter.

5.2 Chapter outline

This chapter presents data on the spread of RSV in the study households. The aim is to ascertain where infants derive their infection from and whether particular members of a household play significant roles in transmission to infants. A subsequent chapter (Chapter 7) will address in details the estimation of transmission probability of RSV in households and in the community.

5.3 Methods

The study methods were described in Chapter 3.

5.3.1 Statistical analysis

The set of terms and their definitions used in the data analysis are given in Tables 3.1, and 4.1. The characteristics of households with infant infections were compared with those with no infant infections. Infected individuals were categorised as (i) study infant, also referred hereafter as the infant, (ii) older child (sibling or cousin), (iii) mother, (iv) father or (v) other adults in the household in line with description provided in methods section of Chapter 4. Infant infections were categorised as originating from outside the household when the infant was the only primary case or was the first individual identified RSV positive in a household outbreak, and from within the household if at least one individual other than the study infant was the first identified as RSV positive in a household outbreak. The origin was deemed inconclusive where the infant and another household member were concurrently first to be identified as RSV positive (i.e. co-primary cases) during a household outbreak. For within household spread and co-primary cases, RSV attachment (G) gene nucleotide sequences (see method below) for the primary case(s) and infant infections were compared.

5.3.2 Attachment (G) gene sequencing and phylogenetic analysis

Sequencing targeted the long ectodomain region of the RSV G gene as previously described (Agoti *et al.* 2012). The sequences were aligned using the Bioedit program (http://www.mbio.ncsu.edu/bioedit/bioedit), RSV group A and B separately. Comparison of the primary-infant cases for RSV group A involved a 648 nucleotide long region and a 732 long region for RSV group Bs. Only one RSV positive specimen (with the lowest Ct values) from each of the pairs was selected for sequencing. Phylogenetc trees were constructed in MEGA 5 program with Maximum Likelihood methods and branch support was assessed by 1000 bootstrap iterations.

5.4 Results

Details on household recruitment and loss to follow-up are found in the results section of Chapter 4. Of the 47 households with complete follow up data, three households were excluded from analysis because the study infant was not sampled during the periods when RSV was detected in their households. All subsequent analyses in the current chapter thus include data from the remaining 44 households.

5.4.1 Baseline characteristic of the 44 households and their members

The baseline characteristics are similar to those reported in the results section of Chapter 4. In summary, the household mean occupancy was 10 members with a range of 4 to 37 members, Table 5.1. The prevalence of males per household was on average 46.5 (95% confidence Interval, CI, 41.2 – 51.8) while the median (IQR) number of older children per household was 4 (3 – 6). Members in each household were on average aged 15.9 (range, 9.3 – 24.6) years. The median (IQR) age of the 44 study infants was 4.2 (2.4, 6.4) months and range of 13 days to 9.9 months, at recruitment. Twenty-one (47.7%) were male (Table 5.1).

Table 5.1: Baseline characteristics of the study participants and households

Characteristic		Description	Statistics
a) Study infants (N=44)			
Male gender		Number (%)	21 (47.7)
Age at recruitment in months		Median (IQR)	4.2 (2. 4 – 6.4)
b) Other household members (N=407)			
Male gender		Number (%)	186 (45.7)
In school		Number (%)	154 (37.8)
Age in years at recruitment		Mean (IQR)	12.6 (6.6 – 26.5)
Frequency distribution	0-<1y	Number (%)	6 (1.5)
of members by age in	1-4y		74 (18.2)
years at recruitment	5-14y		153 (37.6)
	15-39y		133 (32.7)
	40+y		41 (10.1)
c) Households (N=44)			
Household sizes		Median (IQR)	8 (6 – 11.5)
No. of older children (1-14 years) per HH		Median (IQR)	4 (3 – 6)
No of school-going children per HH		Median (IQR)	4 (3 – 6)
No. of members living in the same house as the study infant		Median (IQR)	5 (3 – 6)
Mothers with no formal education		Number (%)	10/43 (23.3)
Duration of HH follow-up in weeks		Median (IQR)	24.8 (23.6-25.3)

Key: HH, household; IQR, interquartile range

5.4.2 Household visits and sample collections

From the 44 households, a total of 16434 contacts with the participants were made obtaining 15,396 (93.7%) NPS. The median (IQR) number of NPS collections per individual was 39 (30 – 42). Further details on the overall sample collections are as found in Chapter 3.

5.4.3 RSV infections in households

Overall, 84.1% (37/44) and 38.4% (173/451) of the households and individual participants, respectively, had one or more episodes of RSV infection. The RSV infection patterns are presented in Appendix T. A total of 73 separate introductions into households were identified with 32 (43.8%) resulting in a household outbreak. The 32 outbreaks were identified in 27 households: one outbreak in 23 households (8 with RSV group A, 14 with RSV group B and 2 with a co-infection of both) and two outbreaks in 4 households (2 with RSV group A followed by group B, and 2 with consecutive RSV group B). The study infant was infected in 28 (87.5%) of the household outbreaks. 5.2. The age distribution of the RSV infections in the 28 infected infants is shown in Table 5.2. Households with RSV spread among the members had a higher mean number of older children than those without RSV infection, though not statistically significant (mean, 95% CI, of 6.0, 4.4 – 7.6, vs. 3.3, 1.8 – 4.8; P value= 0.08), Table 5.3.

Table 5.2: RSV infections among the 44 study infants

Age in months	No. in study	No. Infected	Crude attack rates
0	1	0	13/19 (68.4%)
1	6	4	
2	6	5	
3	6	4	
4	5	3	9/16 (56.3%)
5	5	3	
6	6	3	
7	5	2	6/9 (66.7%)
8	1	1	
9	3	3	
Total	44	28	63.6%

Key: 1, age at the start of sampling by month

Table 5.3: Characteristics of households with and without RSV infection episodes

Characteristics[1]	Uninfected HHs (n=7)	Infected HHs with no virus spread (n=10)	Infected HHs with virus spread (n=27)[2]
Household size	8 (5– 9)	8 (7 – 9)	10 (6 –15.5)
Mean age in years per HH	16.3 (13.4 – 19.2)	15.5 (13.0 –19.0)	14.6 (13.5 –17.5)
No. of older children in HH	3 (2 – 4)	4 (4 –5)	5 (3 – 8.5)
Male: female ratio	1.7 (1.1)	1.0 (0.5)	1.5 (1.2)
No. of school going children	3 (2 – 4)	4 (3 – 5)	4.5 (3 –7.5)
No. of children living in same house per HH	5 (3 – 6)	6 (5 – 7)	4 (2.5 – 6)
No. of children sleeping in same room per HH	3 (2 – 4)	3 (3 – 5)	3 (2 – 3.5)
No. of children sleeping in same bed per HH	2 (2 -2)	2 (2 – 2)	2 (2 – 2)

Key: HH, household; 1, Reported statistics are the median and interquartile range except for age which is mean and 95% confidence interval; 2, limiting this column to the 24 households with outbreaks involving the study infant does not significantly change the statistics shown

5.4.4 Characteristics of households with and without study infant RSV infections

There was no statistically significant difference in households with study infant infections relative to those without in terms household size, household composition and the infant characteristics, Table 5.4.

Table 5.4: Characteristics of households with and without study infant RSV infections

Characteristic[1]	HHs with no infant infection (n=16)	HHs with infant infection (n=28)	P value
Household size	8 (7 – 9)	9 (6 –14.5)	0.4
Mean age of HH members, years	16.0 (14.2 – 17.8)	15.8 (14.3 – 17.2)	0.8
Age of the study infant, months	4.3 (3.1 – 5.5)	4.7 (3.7 – 5.7)	0.6
Number of older children in HH	4 (3 – 5)	5(3 – 7.5)	0.2
Percentage Male per HH	50.0 (40.0 – 57.1)	41.4 (35.8 – 57.7)	0.4
No. of school-going children	4 (2 – 4)	4 (3 – 7)	0.2
No. of HH members living in same house with study infant	6 (4.5 – 6)	4.5 (3 – 6.5)	0.2

Key: HH, household; 1, Reported statistics are the median and interquartile range except for age, mean and 95% confidence interval

5.4.5 Who introduces RSV into the households

From the 73 household episodes, a total of 71 different individuals were primary cases among which 31 (43.7%) individuals were associated with the 32 household outbreaks detected, Table 5.5. Primary cases of all household episodes were study infants, 15 (20.5%); older children, 36 (49.3%); study infants and older child, 4 (5.5%); mothers, 9 (12.3%); fathers, 4 (5.5%); and other household members, 5

(6.8%). The corresponding distribution of the primary cases for household outbreaks was 7 (21.9%), 16 (50.0%), 4 (12.5%), 2(6.3%), 1(3.1%), and 2 (6.3%), Table 5.5.

Table 5.5: Characteristics of the primary cases of the RSV infections in the household

	HH episodes[1] (N=41)		HH outbreaks[2] (N=32)		HH outbreaks with infant infection[3] (N=28)	
Characteristic	*n*	*%*	*n*	*%*	*n*	*%*
Relation to the study infant						
Self	8	19.5	7	21.9	7	25.0
Self and sibling[4]	-	-	4	12.5	4	14.3
Sibling or cousin[5]	20	48.8	16	50.0	12	42.9
Mother	7	17.1	2	6.3	2	7.1
Father	3	7.3	1	3.1	1	3.6
Others	3	7.3	2	6.3	2	7.1
Age groups in years						
<1 year	11	26.8	10	31.3	10	35.7
1– 4 years	5	12.2	6	18.8	6	21.4

5 – 14 years	13	31.7	11	34.4	7	25.0
15 – 39 years	10	24.4	4	12.5	4	14.3
≥40 years	2	4.9	1	3.1	1	3.6
Male gender	18	43.9	17	53.1	15	53.6
Older children	18	43.9	17	53.1	13	46.4
Older children in school	14	82.5	15	88.2	15	84.6

Key: HH – household; 1, limited to household episodes with no spread within the household; 2, includes all household outbreaks, four households had two outbreaks detected; 3, includes household outbreaks associated with the study infant infection; 4, co-primary cases of the study infant (self) and siblings; 5, either sibling or cousins to the study infant

5.4.6 Source of the infant infections

Of the 73 separate household episodes identified, 37 were linked with study infant infections involving 28 of the 44 (63.6%) study infants: 21 had one episode, 5 had two episodes and 2 had three episodes. Household episodes of infection that resulted in the infection of the study infant are depicted in Figure 5.1. The 28 first infant infections were associated with an outbreak in 24 (85.7%) households. Subsequent results on the infant infections were limited to the first episodes, when the infants were RSV naïve. Seven infants (25.0%) had RSV A infection (s), 15 (53.6%) RSV B infection(s) and the remaining 6 (21.4%) had both RSV A and B infection in their first episode. Based on the temporal pattern of infections in each household, 15 (53.6%) of the 28 study infant infections were acquired through transmission within household while 9 (32.1%) were acquired from outside the household. The source of infant infections in the remaining 4 (14.3%) households was inconclusive because the household episodes had a sibling and the infant as co-primary cases, Figure 5.2. Samples were successfully sequenced for the RSV G gene region in 10 of 15 primary case and study infant pairs with all pairs showing complete homology of nucleotide sequence (Table 5.6 and Figure 5.3.) In the remaining 5 in which sequencing failed one or more samples of each pair had a high cycle threshold (Ct) value from the real time M-PCR, a proxy for low virus load. Sequences from the co-primary cases were identical in three of the four sibling-infant pairs. The primary cases for the 15 study infants who acquired the infection from within the household were older children (11), mothers (1), fathers (1) and other adults (2). A similar distribution of primary cases was seen for analysis of all household episodes and outbreaks (Table 5.4). Ten of the 11 older children who were primary cases were attending school and one, an

adult (uncle to the study infant, aged 17 years), was in secondary school. The mean age at recruitment of these older children was 6.9 years with the youngest 2.3 years and oldest 11.5 years. In the 4 households where siblings were co-primary ca`ses with infants, two were attending school.

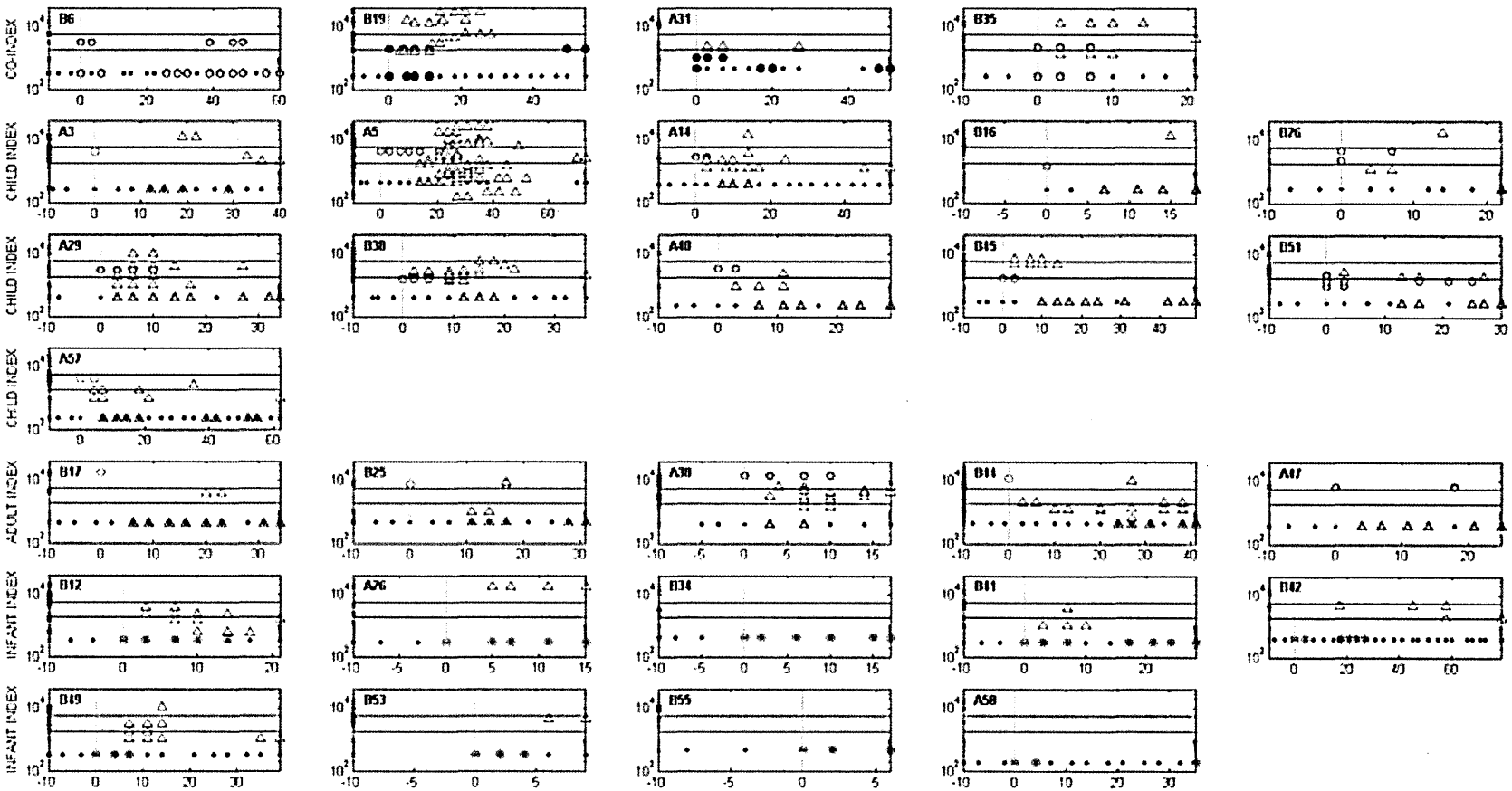


Figure 5.1: Household episodes linked with infant infections. Each panel shows a household episode. The RSV group (A or B) and household number are in the top left of each panel. The vertical axis is the log10 (age) and each individual in the household is shown with a blue cross. Two

horizontal lines show the ages of 5years and 15years. The horizontal axis shows the number days relative to the first (primary) case. Each symbol represents a NPS. For the study infant each NPS is marked with a black dot, but for other members, negative samples are omitted for clarity. Circles show the index case and the top row show households with co-primary cases. Non-primary cases are shown with triangles. Filled symbols show that the virus from this infection was successfully sequenced. If the colour is black, then the viruses are different, and if the colour is green, the viruses are the same

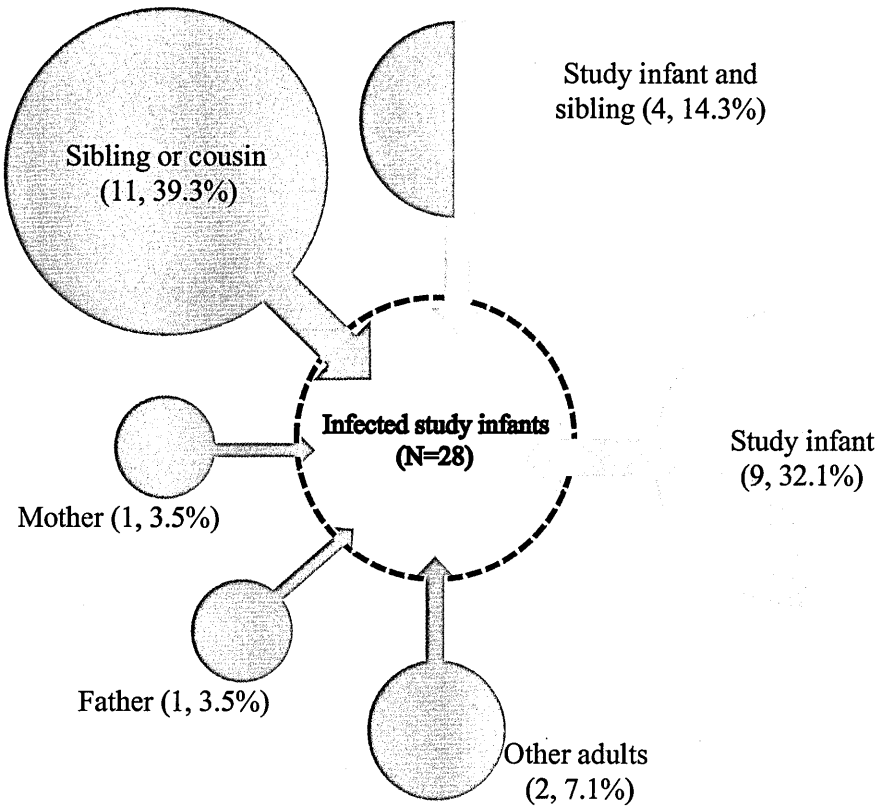


Figure 5.2: Distribution of the primary cases linked with the 28 study infant infections in rural Kenyan households. Only the first household episodes/outbreaks involving the study infants were considered. The light and dark shades indicate outside and within household acquisition of the infant infections, respectively, and the area of the circle is in proportion to numbers of cases in each category

Table 5.6: Sequencing the G gene of infant-primary case pairs

HH ID.	Relation[1]	Sequenced[2] *Infant?*	Sequenced[2] *Index?*	G gene	No. of nt changes[3]
26	Sibling[4]	Yes	Yes	Identical	-
51	Sibling[4]	Yes	Yes	Identical	-
16	Sibling	Yes	Yes	Identical	-
29	Sibling	Yes	Yes	Identical	-
40	Sibling	Yes	Yes	Identical	-
57	Sibling	Yes	Yes	Identical	-
14	Sibling	Yes	Yes	Identical	-
5	Sibling	Yes	Yes	Identical	
38	Sibling	Yes	Yes	Identical	
45	Sibling	Yes	Failed	-	-
3	Cousin	Yes	Failed	-	-
25	Mother	Yes	Failed	-	-
17	Father	Yes	Failed	-	-
44	Other	Yes	Failed	-	-
47	Other	Yes	Yes	Identical	-
19	Self & Sibling	Yes	Yes	Different	1
31	Self & Sibling	Yes	Yes	Identical[5]	-
35	Self & Sibling	Yes	Yes	Identical	-
6	Self & Sibling	Yes	Yes	Identical	-

Key: 1, relation of the primary cases to the study infant; 2, the G gene region sequenced in RSV positive samples where the study infant appear to have acquired the infection from within the household or the study infant was a co-primary case; 3, all nucleotide changes were non-synonymous; 4, more than one sibling was primary cases; 5, two household

participants were not sampled while the infant got the infection and one was shedding the virus on the first sample collected and the virus in the infant matches with virus in the index but one position showed mixed population; HH ID, household identifier; nt, nucleotide ; index refers to the primary and co-primary case

a) RSV group A (b) RSV group B

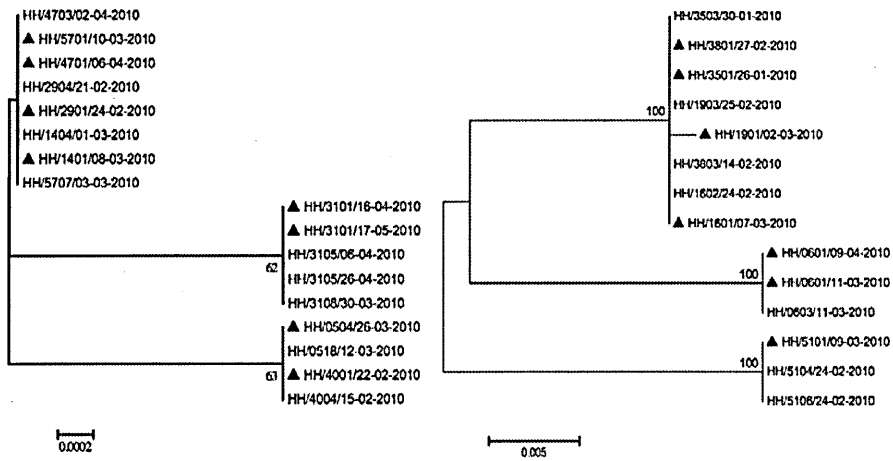


Figure 5.3: Phylogenetic tree showing the G gene similarity in the study infant-primary case pairs for RSV (a) group A and (b) group B. Infant cases are preceded by a dark filled triangle. The scales for branch lengths of group A versus B are different. The name of the taxon HH/ household number with individual ID/ date of sampling

5.4.7 Timing of RSV infections among the study infants

When considering day 0 to be the day of first RSV detection in the household outbreaks, the median (IQR) duration to infection of the study infant, first older child, mother, father and the first other household adults was 5 (0.8 – 5), 0(0 – 3), 5 (3 – 12), 18 (8.5 – 21.5) and 14.5 (0 – 15) days. Overall the infants were infected early in course of household outbreaks, Figure 5.4.

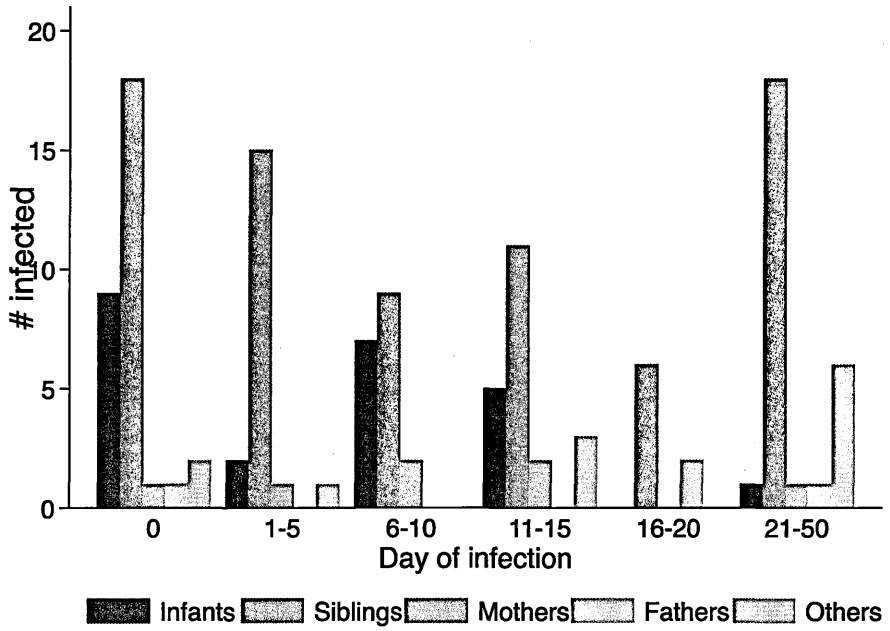


Figure 5.4: Distribution of days to onset of RSV shedding during household outbreaks by household members. Day 0 is the first day of RSV detection in the household

5.5 Discussion

The Chapter reports on analysis designed to determine who introduces the RSV infection into the household with particular emphasis on the source of RSV infection to RSV naïve infants, as those who are most vulnerable to disease from RSV infection. The aim was to establish the proportion of infants infected from a source from within the immediate household and, if so, which household member was responsible for introducing RSV into the infant household. Data from rural coastal Kenya are presented for 44 households with 451 members recruited and followed actively for NPS collection every 3-4 days regardless of symptoms during a full RSV epidemic. RSV infection was detected in 84% of the households and 38.4% of the participants. Both the two known RSV subtypes (A and B) circulated in the community during the study period, which assisted in distinguishing transmission events.

The study infant was infected at least once in 28 of the 44 households with about a third (32.1%) acquiring the infection from outside the household and about half (53.6%) from within the household. In the remaining (14.3%) households, the source of infant infection was undetermined as the infant and sibling were co-primary cases. Children attending school were key (80%) contributors in the within household transmission. Mothers or fathers were less likely to be the primary cases. There were no statistical differences between the individual or household characteristics of infected and uninfected infants, suggesting that differences between households are largely based on chance or unobserved characteristics. In addition, the distribution of the primary cases for all household episodes and outbreaks of RSV was similar, indicating it is not the person who brings the infection into the household that determines if the infection spreads or not, but may be chance, or an unobserved characteristic of the household such as whether an outbreak was experienced in the previous year(s).

Only one study in the past has looked at the spread of RSV within households in detail (Hall *et al.* 1976). Hall and colleagues recruited 36 US families with young infants, collecting nose or throat samples twice-a-week regardless of symptoms during two months of an RSV season, assessing infection using viral culture (Hall *et al.* 1976). RSV was detected in 16 (44.4%) of the families and 21.9% of the 188 members, which were lower than in the current study (84.1% and 38.4%, respectively). Based only on temporal occurrence of infections, RSV appeared to be introduced into the 16 infected families by siblings in 8 (50%) families; other household members including parents in 3 (18.8%) families and 5 (31.3%) involved an infant (co-primary cases). The corresponding statistics for the 32 household outbreaks recorded in our study, was 16 (50%), 5(18.8%) and 11 (34.4%). As in our study, these findings highlight the importance of older children in introducing RSV into households with young infants. Our study also shows that children with daily contact with many other children

particularly in schools are important in bringing RSV infections into the households. Other less comparable studies have implicated siblings (Badger *et al.* 1953; Berglund 1967; Fox *et al.* 1975; Okiro 2007; Crowcroft *et al.* 2008) and mothers (Stensballe *et al.* 2004) with infection of the infant in the family. However, none of these previous studies (Badger *et al.* 1953; Berglund 1967; Fox *et al.* 1975; Stensballe *et al.* 2004; Okiro 2007; Crowcroft *et al.* 2008) were designed specifically to identify the source of infant RSV infections in families. The present study and that of Rochester family study (Hall *et al.* 1976) were designed for the purpose of identifying transmission chains by sampling frequently and irrespective of symptoms. However, methodologies have changed since that of Hall et al (Hall *et al.* 1976). The current study has the advantages of increased sensitivity in case detection using PCR and of sequence data support for observed primary-infant infection pairs. Additionally, the US study is likely to have missed infections early in the RSV epidemic due to sampling delay and suffered a high proportion of co-primary cases restricting the ability to identify the relative individual contributions as infection sources into the families. Nevertheless, the consistency of results given the contrasting locations and different times, suggest that the important role of a school-age sibling in bringing infection into the household is not unique to a particular social and demographic setting.

The results support the notion that preventing infection in school-going children could indirectly reduce RSV infection in infants. However, an assessment of the impact of such an intervention requires consideration of the competing risks, e.g. somebody else might introduce infection if not the vaccinated individual, or the infant might be infected in the community if not the household. However, targeted sibling immunisation would nonetheless result in fewer susceptible individuals within the household which would reduce spread and provide indirect infant protection, and any reduction in rate of infection to infants will translate into a delay in infection to an older age which is associated with reduced risk of

severe disease. These results also point to schoolchildren as the "core-group" for RSV. Consequently, universal immunisation of children (regardless of their sibling status) could have an impact on circulation of RSV in the community (especially in communities such as this one in typical rural Africa with a large numbers of children), thus further reducing the risk of infant infection. However, such an intervention is more likely to interrupt the transmission dynamics of RSV and potentially change seasonality and the age-distribution of susceptibility and infection. A postdoctoral scientist in our research group in Kilifi, is currently conducting modeling studies to further understand the detailed transmission within the household and the impact of mass immunization to address these issues.

The current study has also some limitations. First, generalizability of the results is limited in that households were specifically selected if they had an older child, and the number of households was small in number and within a tight geographical area populated by rural farmers. Second, the frequency of sampling might have missed short duration RSV shedding (if less than 3 days), including primary cases, despite the intensive sampling. For some cases where the infant is the primary or co-primary, the index case may have gone unidentified. This risk is bigger than the alternative (where the infant was actually the first but was not detected), because the sampling of infants was more complete and was mainly infected in the first few days once the infection was introduced into household. It is therefore plausible that our estimate of the importance of within household spread as a source of infant infections is an under-estimate. Third, given that diagnosis of RSV in adults is more difficult than in children (Falsey *et al.* 2003), the role of parents or other adults in household RSV may have been under-estimated. However, the use of sensitive real time PCR will have mitigated this to a significant degree. The use of serological identification of cases might have reduced misclassification of infection but would not have assisted in resolution of transmission chains. Following a recent collaboration with Sanger Institute in United Kingdom, there is

ongoing work using next generation sequencing and full genome sequencing to identify any molecular fingerprinting that might help to delineate the chains of RSV spread within the households.

In conclusion, the study demonstrates that in this rural Kenyan location, for a larger proportion of infant RSV infections the most likely source is from an individual from within the household and also predominantly introduced into the household by an elder school-going sibling. Thus there is potential for targeted immunisation of older siblings of naïve infants, or universal vaccination of older infants and children, in reducing the spread of RSV and the risk of infection entering households and delay first infection in infants to an older age where associated severe disease is less of a risk.

CHAPTER SIX

6 Duration Of RSV Shedding

6.1 Introduction

RSV primarily infects the upper respiratory tract, with predilection for the deep nasopharynx, where it multiplies and is shed from epithelial cells to exit via nasal secretions. The virus also spreads to the lower respiratory tract infection (LRTI) but only in a small proportion of infected individuals, suggesting that upper respiratory tract infection (URTI) is more important than LRTI in the shedding and spread of virus, although this will also depend on the viral load shed. Community-based studies are thus likely to provide a more complete representation of the RSV shedding patterns. Further more, RSV repeatedly infects throughout life (Henderson *et al.* 1979; Glezen *et al.* 1986; Agoti *et al.* 2012). While repeat infections are considered less severe relative to the primary infections, they constitute the bulk of RSV infections during the annual epidemics observed worldwide. The role of the repeat infections in the overall transmission of RSV in the population is likely to be related to the amount and duration of shedding.

Studies estimating duration of RSV shedding reviewed in Chapter 2 (Table 2.3) report the duration to be in order of days albeit with considerable variations in the estimates. The variation in the reported estimates could be explained by differences in: (i) sampling regimes i.e. the type of specimen and frequency of collection and whether collected regardless of respiratory symptoms; (ii) differences in study settings: hospital studies are biased to young children with severe RSV disease with the additional problem of failing to objectively establish the start, and, in some instances, the end of shedding particularly when patients recover before they stop shedding, a phenomenon in RSV (Frank *et al.* 1981) and, (iii) differences in sensitivity of the specimen screening methods, culture versus

immunofluorescence methods. Early studies relied much on viral culture and with labile nature of the RSV virus specimen handling was likely to affect the sensitivity especially if samples were not inoculated immediately on collection, a concern raised in some of the studies (Frank *et al.* 1981). The current study was designed with the above factors in consideration. Given that samples were collected every 3 to 4 days regardless of symptoms and tested using PCR, the accruing RSV infection data is detailed to allow estimation of shedding duration. Even with this design, the problems of ascertainment and analysis (i.e. censoring and test sensitivity) cannot be completely eliminated. Short infectious episodes (occurring between sampling times) would be missed.

Other viruses infecting the respiratory epithelium might interfere with RSV infections and possibly alter the disease severity if they occur concomitantly. Presence of other respiratory viruses, before or during RSV infections, might alter the RSV infection duration since they are likely to have ecological competition. These possibilities have received very little attention in the literature, and are explored in the current Chapter.

Estimates of the duration of shedding are key to delineating the virus transmission dynamics in mathematical models to predict the potential population impact of vaccines against RSV. The basic reproductive number is directly proportional to the duration of viral shedding, for instance the basic reproductive number would be greater by two times if the duration of shedding was doubled in an randomly mixing and susceptible population, underscoring the importance of realistic estimates. In addition, development of RSV control strategies is likely to depend on the mechanisms of the virus persistence and spread within a population, which are intrinsically related to the viral shedding patterns. Detailed data on the shedding duration among non-hospitalized individuals in relation to factors such as age, RSV group (A or B), infection severity, presence of viral co-infections and gender would help in identifying from whom most viruses come from.

6.2 Chapter outline

In this chapter, RSV shedding duration is estimated from all episodes of RSV infections within the household study. The effect of age, RSV group, clinical severity, gender, presence of concurrent infections and other persons infected in the household on rate of cessation of RSV shedding is assessed. Of these, the factor of key interest is age, which can also be used as a proxy for the number of previous infections.

6.3 Materials and methods

The data arise from the household study, which was described in Chapter 3. In this Chapter, data from the 179 individuals from 40 households infected with PCR positive NPS during the study period were included.

6.3.1 Statistical analysis

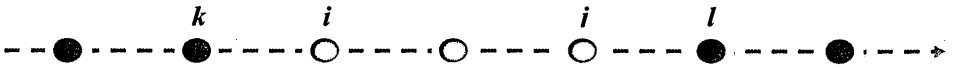
Data were analyzed using STATA version 11.2 (Statacorp, College Station, Texas, US) and first converted to time-to-event data. Student's t-test, Mann–Whitney U test, chi-squared test, and Fisher's exact test were used as appropriate for descriptive analysis. The time of sample collection was recorded and plotted graphically to determine if the sensitivity of the test (i.e. viral shedding) had an obvious diurnal rhythm or collection bias. An individual RSV episode was defined as the period within which an individual provided specimens which were PCR positive for the same infecting RSV group with no more than 14 days separating any two positive samples (defined in Table 3.1). Episodes whose first sample of the infection was positive for both RSV A and B were considered as a co-infection and counted as one individual infection episode.

6.3.2 Estimation of shedding durations

Since the exact date of start and end of RSV shedding could not be directly determined from the interval censored data, three estimates of shedding duration were made based on the

following criteria as illustrated in Figure 6.1. Let the first day of an infection episode be defined as *i* and the last day as *j*, and the first negative sample day prior to the infection episode defined as *k* and the first negative sample after the last positive sample of an episode defined as *l*, assigning day *k* arbitrarily a value 0 increasing by daily units up to day *l*. Further, assume that *i* and *j* contribute each one full day of shedding and days *k* and *l* contribute 0 days of shedding, then for:

a) Minimum estimates: shedding begins on day *i* and ends on day *j* (inclusive) with the shedding duration defined as $j-i+1$ days.
b) Midpoint estimates: shedding begins mid-way between day $k+1$ and day *i*, and ends mid-way between date *j* and date $l-1$, with the shedding interval defined as $(j+(l-j)/2)-(i-(i-k)/2)$. For episodes, whose interval between the last negative and first positive (left censored) and last positive and first negative (right censored) was greater than 7 days, one half of the average interval between sample collection was subtracted from the date of the first positive and added to the date of the last positive sample, respectively, to define the start and end of shedding.
c) Maximum estimates: shedding begins a day after k and ends day before l, hence the shedding duartion is defined as $l-k-1$. For left and right-censored episodes the average interval between sample collections was subtracted and added to define start and end of the shedding respectively.



Key: ● Negative sample ○ Positive sample

Figure 6.1: Schematic diagram illustrating the timeline for each individual and important time points used in estimating shedding durations. Minimum estimates, j–i+1, Midpoint estimates, (j+(l-j)/2)-(i-(i-k)/2), and maximum estimates, l-k-1

The person-time was calculated as the number of days from start date of shedding to the end of the infection episode (recovery date). An individual whose last NPS was PCR positive or the last sample of an infection episode was followed by duration of more than 7 days without NPS collection was considered right censored. Recovery rates were calculated as the number of recovery events, in RSV infected individuals, divided by the sum of person-days of being PCR positive. Kaplan-Meier plots were used to show the distribution of survival times (probability of shedding cessation over the observed PCR positive periods) for all categorical variables explored. Estimation of shedding durations (and recovery rates) by age, gender, presence of acute respiratory infection (ARI) symptoms, and presence of other viral co-infections were made. ARI was defined as presence of one or more of the following respiratory symptoms at the time of sample collection; cough, nasal discharge/blocked nose, or difficulty in breathing (Table 3.1). Co-infection was assigned when within the RSV infection episode any sample was PCR positive for another virus (see Table 4.1 for definition). Presence of other respiratory viruses in the samples collected in the period of 14 days prior to start of RSV episodes was also defined. Household outbreak was defined as a period within which one or more individual episodes occurred in members of the same household without an interval of 14 days or more in which a PCR positive specimen was absent from the household. The proportion of the household members infected during household outbreaks measured the intensity of the outbreak.

6.3.3 Cox regression analysis

Cox proportional hazards model was used to assess factors influencing the RSV shedding durations. The null hypothesis was that there is no association between the duration of viral shedding and age, presence of ARI or other viral infections or concurrent infections in the household. The unadjusted hazard ratios (HR) from the univariate analysis based on the data from the three approaches are presented but only the midpoint data were used in the final multivariate model. The multivariate Cox regression model was developed by a forward stepwise procedure rejecting variables with a p-value ≥ 0.05 in likelihood ratio tests. Infecting RSV group, gender and detection of a second RSV infection were selected apriori for inclusion in the final model as possible confounders. Covariates were introduced in descending order of strength of association determined from univariate analysis. Robust variance estimator (Huber-White sandwich estimator) was used to derive adjusted 95% confidence intervals (CI) to account for within-household correlation of recovery. The effect of left censoring was accounted for by including a dummy variable denoting the left censored infection episodes or by excluding the left censored episodes. Schoenfeld residuals were used to test the proportionality assumption, which underlies the Cox regression method.

6.3.4 Attachment (G) gene sequencing and phylogenetic analysis

This was aimed at delineating cases of suspected repeat infections by sequencing the part of the G gene and phylogenetic analysis as described in Chapter 3 and 5. For sequencing, only one RSV positive sample (with the lowest Ct values) from each of the paired episodes was selected for sequencing.

6.4 Results

6.4.1 Baseline characteristics of the RSV infected individuals

Data from the household prospective cohort study was used in this analysis with focus on the 179 individuals infected with RSV (Table 6.1). Thirty-one (17.3%) were under 1 year,

41(22.9%) between 1 to 4 years, 66 (36.9%) 5 to 15 years and the rest (22.9%) above 15 years of age at the time of infection. The individuals were between the ages of 1.9 months and 78 years at time of infection; only 9 were above 40 years. The median age (inter quartile range, IQR) was 6.5 (2.4 – 14.5) years, and females numbered 96 (53.6%). A total of 205 infection episodes were recorded with 155 individuals experiencing one episode, 22 with two episodes and two individuals experiencing three episodes during the single RSV epidemic. RSV group A was associated with 88 infection episodes, RSV group B with 113 while 7 episodes were with co-infection of RSV group A and B. Figure 6.2 shows all the observed infection episodes stratified by the infecting RSV group.

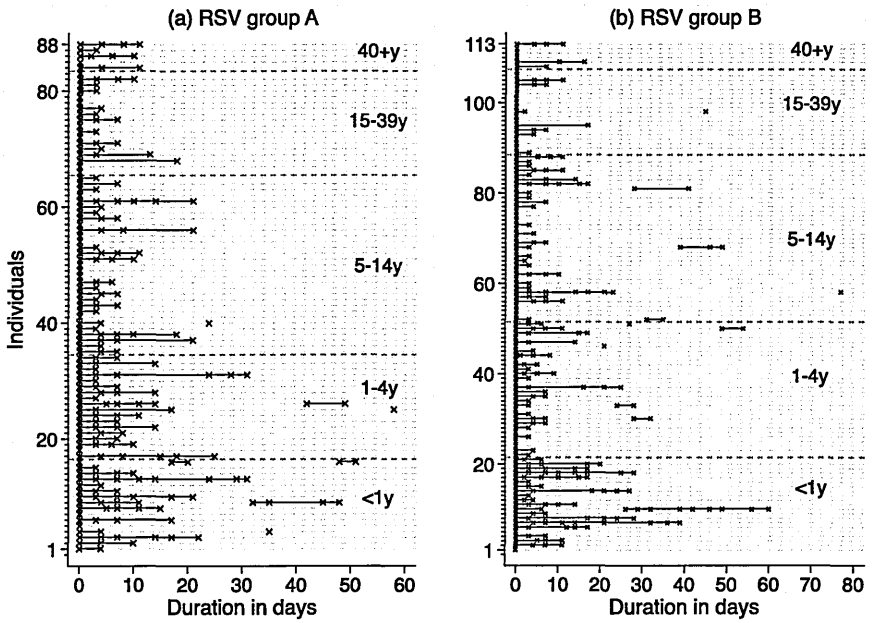


Figure 6.2: RSV group A (a) and group B (b) episodes ordered by age of the individuals infected. RSV positive samples are marked by blue 'x' while the red line links PCR positive samples from the same infection episode. Appendix U shows the RSV infection episodes by order of occurrence within the households.

Table 6.1: Baseline characteristics of the 179 study participants and the associated RSV infection episodes

		Persons[1] (N=493)		Episodes[2] (N=205)		Shedding periods[3] (N=2176.9)	
Covariates	**Categories**	*n*	*%*	*n*	*%*	*n*	*%*
Age in years, at start of sampling	<1y	55	11.2	37	18.1	613.7	28.2
	1-<5y	82	16.6	49	23.9	540.7	24.8
	5-<15y	165	33.5	74	36.1	654.1	30.1
	15-<40y	147	29.8	36	17.6	276.5	12.7
	≥40y	44	8.9	9	4.4	91.9	4.2
Relations[4]	Study infant	47	9.5	34	16.6	586.7	27.0
	Sibling	164	33.3	79	38.5	739.7	34.0
	Cousin	124	25.2	55	26.8	553.1	25.4
	Mother	46	9.3	14	6.8	124.9	5.7
	Father	33	6.7	6	2.9	34.0	1.5
	Other	79	16.0	17	8.3	138.5	6.4
Gender	Female	272	55.2	116	56.6	1223.4	56.2
	Male	221	44.8	89	43.4	953.5	43.8
RSV group	A	-	-	81	39.5	796.4	36.6
	B	-	-	110	53.7	1138.3	52.3
	A and B	-	-	14	6.8	242.2	11.1
Infection	First	-	-	179	87.3	1932.9	88.8
	Second	-	-	26	12.7	244.0	11.2
ARI	No	-	-	87	42.4	635.3	29.2
	Yes	-	-	118	57.6	1541.6	70.8

Coinfection	No	-	-	119	58.1	958.6	44.0
	Yes	-	-	86	42.0	1218.3	56.0
During HH	No	-	-	42	20.5	290.8	13.4
outbreak	Yes	-	-	163	79.5	1886.1	86.6
Outcomes							
Durations[5]	<7 days	-	-	74	36.1	328.0	15.1
	7 – <14 days	-	-	75	36.6	683.9	31.4
	15 – <21 days	-	-	32	15.6	517.5	23.8
	>=21 days	-	-	24	11.7	647.5	29.7

Key: 1, number of individuals in the parent study; 2, RSV infection episodes; 3, total infection durations based on midpoint estimation; 4, the relationship of a household members with the study infant; 5, duration of RSV shedding episodes; ARI, acute respiratory illness; HH, household

6.4.2 Sampling collection times

There was no association of sample collection hour and RSV detection (Figure 6.3), hence all the shedding duration estimates assumed the samples were collected at the same time of the day.

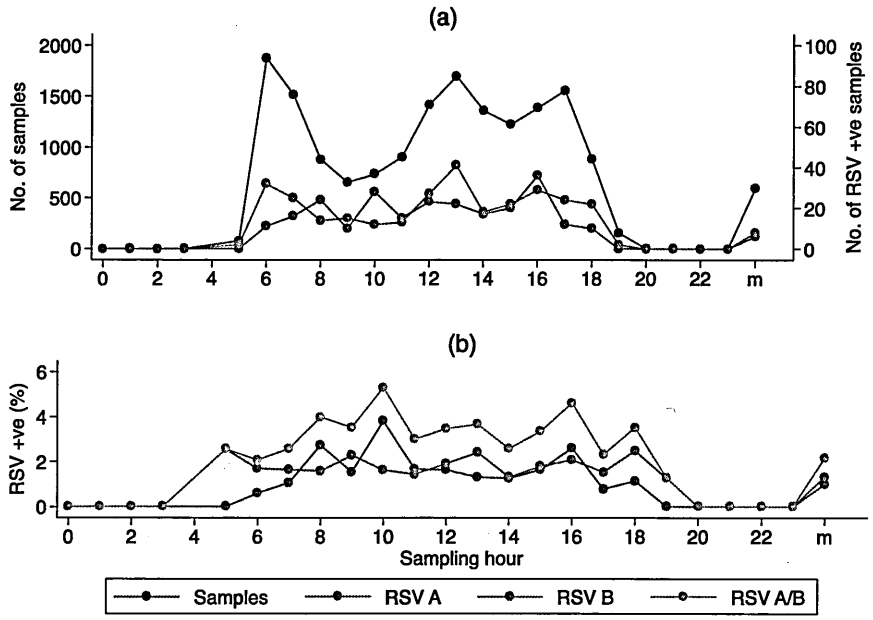


Figure 6.3: (a) Frequency distribution of the total, RSV group A and RSV group B positive, sample collections by time of collection and (b) percentage of the samples positive for RSV group A, RSV group B or either group by time of sample collection. 'm' on the x-axis represents samples with missing time of collection

6.4.3 Censoring and sample collection intervals

Of the 205 RSV episodes, 177 (86.3%) were fully observed i.e. negative samples before and after the RSV episode were collected within 7 days. Two individuals tested positive in their first sample and 9 infection episodes were detected following a period of absence of more than 7 days from the household (left censoring). Right censoring was observed in 15 infection episodes; in 4 episodes the last PCR positive sample was the last sample collected for the individual while in 11 the last positive sample was followed by a period of no sample collection (mainly occurring when the individual travelled) for more than 7 days. Two episodes had both left and right censoring. Table 6.2 shows the distribution of sampling intervals, age at infection and

estimated shedding duration by the censoring type. The average interval between sample collections during RSV infection episodes was 3.5 (95% CI, 3.4 – 3.6) days with a median (IQR) of 3 (3 – 4) days (Figure 6.4). The mean intervals prior to and immediately after the infection episodes for the fully observed episodes were not statistically different; 3.7 (95% CI, 3.6 – 3.9) versus 3.7 (95% CI, 3.6 – 3.8) days respectively; t-test P-value=0.4819. The correlation of sampling intervals before and after RSV episodes is graphocally shown in Appendix V.

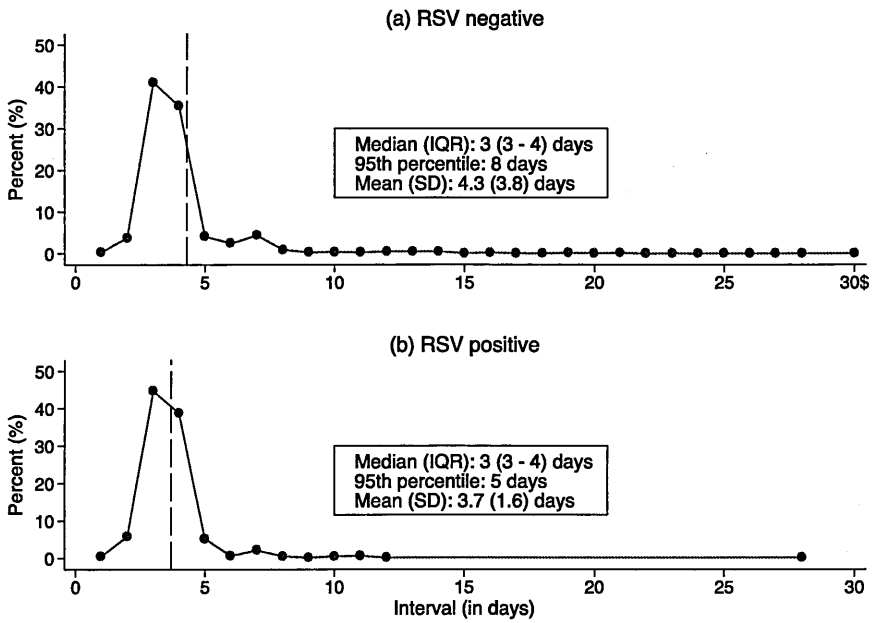


Figure 6.4: Distribution of the intervals between NPS collections during RSV negative (a) and PCR positive periods. The red dashed vertical line represents the mean interval; $, in 29 instances the interval between RSV negative samples was more than 30 days

Table 6.2: Sampling intervals, age at infection and estimated shedding duration by the censoring type.

Characteristic		All[1] (n=205)		Fully[2] (n=177)		Left[3] (n=11)		Right[4] (n=15)		Both[5] (n=2)	
		Mean	*95% CI[6]*	*Mean*	*95% CI*	*Mean*	*95% CI*	*Mean*	*95% CI*	*Mean*	*95% CI*
Intervals[7]	Before	4.2	3.8 – 4.5	3.7	3.6 – 3.9	9.6	8.6 – 10.5	3.4	2.8 – 4.0	18.50	-
	During	3.5	3.4 – 3.6	3.5	3.4 – 3.6	3.8	2.8 – 4.9	4.1	3.1 – 5.0	-	-
	After	5.2	4.2 – 6.3	3.7	3.5 – 3.8	3.4	2.4 – 4.3	24.1	13.3 – 34.9	11.0	-
Age in years[8]		11.1	9.3 – 13.0	10.4	8.6 – 12.3	15.1	5.0 – 25.1	15.6	6.5 – 24.6	16.9	-
Shedding duration[9]		10.6	9.6 – 11.7	10.9	9.7 – 12.1	7.4	4.5 – 10.2	8.9	6.7 – 11.0	14.8	-

Key: 1, all the RSV infection episodes; 2, fully observed episodes; 3, left censored episodes only; 4, right censored episodes only; 5, episodes with both left and right censoring; 6, 95% CI, 95% confidence Interval; 7, intervals between samples before, during, and after the RSV infection episode in days; 8, age in years at the start of the episode; 9, shedding duration in days based on the midpoint approach

6.4.4 *Duration of RSV shedding*

The mean duration (95% CI) of RSV shedding based on minimum, midpoint and maximum estimates was 8.6 (7.5 – 9.7), 11.2 (10.1 – 12.3) and 14.0 (12.8 – 15.2) days (Table in Appendix W). The corresponding mean durations of shedding for the fully 'observed' episodes were 8.2 (7.1 – 9.4), 10.9 (9.8 – 12.1) and 13.6 (12.4 – 14.8) days. The duration of shedding from the midpoint approach had a median (IQR) of 8.5 (5.5 – 14.0) days and range of 1 – 44 days (Figure 6.5). The RSV recovery rates are shown in Appendix X.

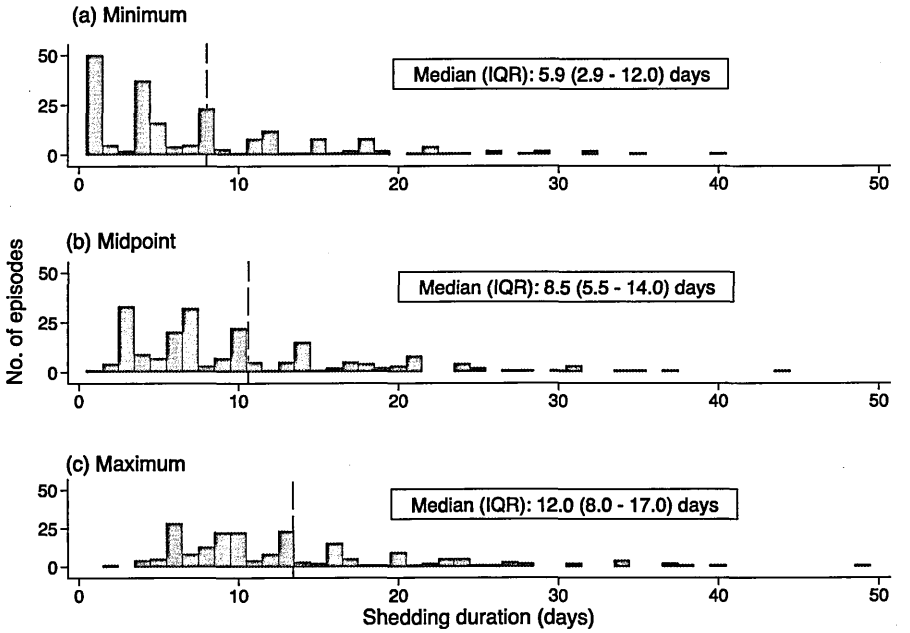


Figure 6.5: The frequency distribution of the days to cessation of shedding by method of estimation using the data from all the 205 episodes. The dotted red line shows the mean duration of shedding (see text for values). IQR, interquartile range

6.4.5 *Univariate analysis: Factors influencing the rate of cessation of RSV shedding*

The Kaplan-Meier plots showing the association of the various factors and recovery probability arising from midpoint analyses are shown in Figures 6.6 and 6.7. The

hazard ratios (HR) for the various factors from univariate Cox regression were similar for minimum, midpoint and maximum estimates data and are shown in Appendix Y.

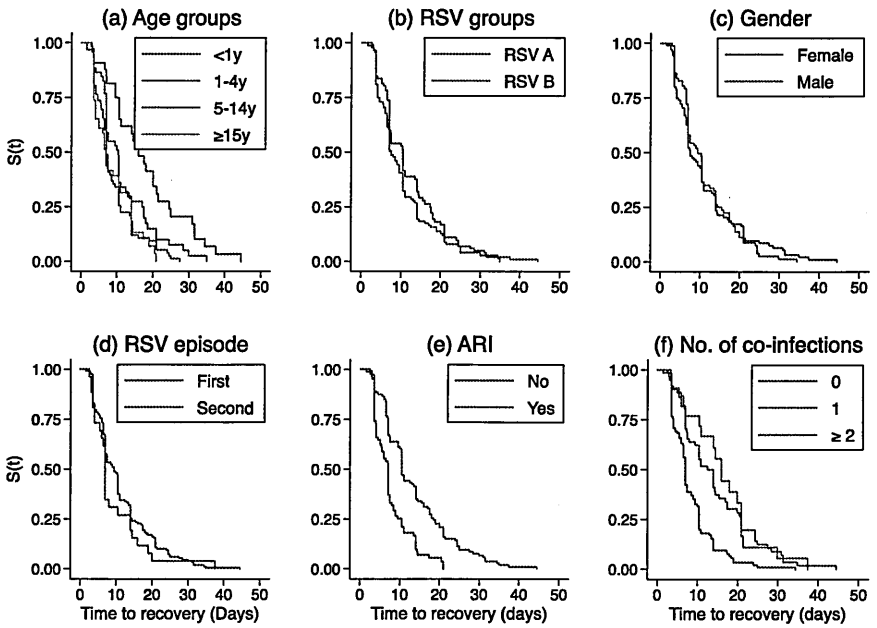


Figure 6.6: Kaplan-Meier survival function plots of midpoint data for cessation of RSV shedding stratified by age at infection in years (a), RSV groups (b), Gender (c), order of RSV infection episodes (d), presence of acute respiratory symptoms during the RSV episode (e) and number of other co-infecting viruses (f). The legend for each graph is shown in the respective inserted text box.

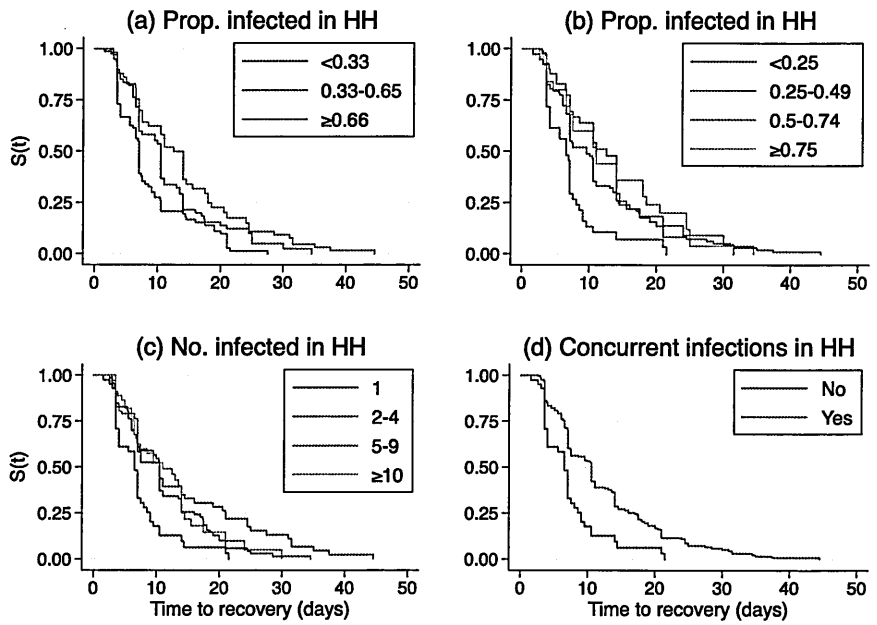


Figure 6.7: Kaplan-Meier survival function plots for cessation of RSV shedding stratified by various markers of concurrent RSV infections within the household (HH). Graphs (a) and (b) show results by the proportion of household members infected; graph (c) by the number of persons infected and graph (d) by the presence of another person infected in the household. The legend for each graph is shown in the respective inserted text box.

6.4.6 Multivariate analysis: Factors influencing rate of RSV recovery

The midpoint data were taken forward for the multivariate analysis and the final Cox regression model is reported in Table 6.3. The results were similar with and without inclusion of the left censored RSV episodes (see Appendix Z for final model with left censored data). The proportionality assumption in the final Cox regression model was not violated based on the test of the Schoenfeld residuals (see Appendix AA).

i) Age of infection

The duration of shedding was significantly associated with age. Infants (<1 year of age) shed the virus for 18.0 (14.2 – 21.8) days while the older individuals shed for shorter durations; 11.8 (9.7 – 14.0), 9.1 (7.7 – 10.4), 8.4 (6.6 – 10.2) and 11.2 (7.4 – 15.0) days in 1 – 4, 5 – 14, 15 – 39 and ≥40 age groups, respectively based on midpoint estimates. Using the infants as the reference group, the adjusted rate of loss of infection was 1.98 (95% CI, 1.30 – 3.02), 1.82 (1.16 – 2.87), 2.10 (1.20 – 3.66) and 1.31 (0.36 – 4.81) times in 1-4, 5-14 years, 15-39 and ≥40 age groups, respectively. Infants aged ≥6 months had a decreased recovery rate by 48% relative to those <6 months (adjusted HR, 0.52; 95% CI, 0.20 – 1.31) but this was not statistically significant (Wald test p= 0.164).

i) Symptomatic infections

The duration of shedding differed between those individuals who had symptomatic infections and those with no symptoms: 13.5 (95% CI, 11.9 – 15.1) vs. 7.8 (6.7 – 8.8) days. The rate of recovery was lower by 44% in symptomatic infections relative to asymptomatic infections (adjusted HR 0.56; 95% CI, 0.40 – 0.79).

ii) Presence of other respiratory viruses

The presence of one or more additional virus (rhinovirus, coronavirus or adenovirus) was detected in 86 RSV infection episodes. Co-infections were associated with increased duration of RSV shedding of 14.8 (12.9 – 16.8) days relative to duration of 8.5 (7.4 – 9.5) when no co-infections were detected. Rate of cessation of RSV shedding was lower by 50% in episodes with co-infection compared to those without (adjusted HR, 0.65; 95% CI, 0.23 – 0.51). Examination of co-detected viruses individually (adenovirus, coronavirus, rhinovirus) or mixed; each identified a significant association with decreased recovery rate relative to RSV infection alone (adjusted HR, 0.40 (95% CI, 0.22 – 0.73), 0.36 (0.21 – 0.62), 0.33 (0.17 – 0.62) and

0.30 (0.17 – 0.50), respectively). This association did not linearly vary with the increase in the total number of the viruses co-detected with RSV. Detection of infection with any one or more of rhinovirus, adenovirus or coronavirus, in the two weeks preceding the start of RSV infection, but not during the RSV episode itself, was associated with 56% increase in the recovery rate of RSV infection (adjusted HR, 1.56 (1.02 – 2.39), p=0.041; Table 6.3). In contrast, RSV episodes associated with detection of other viruses in the two weeks prior to and also during the RSV infection were associated with 52% decrease in rate of recovery relative to those with no other virus prior and during the RSV episode (adjusted HR, 0.48 (0.32 – 0.73); Table 6.3).

iii) Presence of contemporaneous RSV infections in the household

RSV episodes linked with spread of the virus in the household were associated with an increased duration of shedding. The rate of shedding cessation was significantly lower for infection episodes associated with household outbreak relative to single infection episodes (adjusted HR 0.58; 95% CI, 0.43 – 0.78). A variable denoting the proportion of individuals infected during the household outbreak improved the model fit and was used in the multivariate analyses (likelihood ratio test p-value=0.0229), Table 6.3, and Figure 6.7.

iv) Gender

The rate of recovery between males and females did not differ significantly (adjusted HR, 0.97; 95% CI, 0.77 – 1.23), Table 6.3.

v) Infecting RSV group

RSV group B was associated with 9% (adjusted HR 1.07 (0.77 – 1.23) increased rate of recovery relative to RSV group A but this result was not statistically significant (Table 6.3).

vi) Number of episodes

Recovery rate was similar in subsequent RSV infection episodes occurring within the study period compared to the first observed infections episodes (adjusted HR, 0.91 (0.53 – 1.56), Table 6.3).

Table 6.3: Effects of various covariates on recovery rates using midpoint data: Multivariate cox regression analysis[1]

Factors	Categories	Reference	HR	95% CI	P-value
Age groups	1-4y	<1 year	1.98	1.30 – 3.02	0.002
	5-14y		1.82	1.16 – 2.87	0.01
	≥15y		1.97	1.11 – 3.51	0.021
Symptomatic		Asymptomatic	0.56	0.40 – 0.79	0.001
Detection of other[2] viruses before and during RSV episode	Before[3]	No other	1.56	1.02 – 2.39	0.041
	During	viruses	0.48	0.32 – 0.73	<0.001
Other HH members infected (%)	33%	<33%	0.59	0.40 – 0.87	0.008
	66%		0.51	0.35 – 0.74	0
RSV group B		A	1.07	0.83 – 1.39	0.588
Male gender		Female	0.97	0.77 – 1.23	0.804
Second RSV infection		First	0.91	0.53 – 1.56	0.73

Key: HR, hazard ratio; CI, confidence interval; HH, household; 1, left censored episodes excluded; 2, other viruses were adenoviruses, rhinoviruses and coronaviruses; 3, detection of other respiratory viruses during the 14 days prior to the start of RSV episode ONLY

6.4.7 Prolonged shedders

Twenty-four individuals shed RSV for 21 or more days; 10 (41.7%) were < 1 year old, 6 (25.0%) were aged 1-4 years, 7 (29.2%) aged 5-14 years while the remaining was aged 17.2 years. The median (IQR) age at infection was 14.7 (9.4 – 69.7) months with a range of 3.8 months to 17.2 years. Twenty-two (91.7%) of the infection episodes were symptomatic. The prolonged shedders contributed 647.5 days of RSV shedding which was 29.7% of the entire shedding durations from the 205 episodes based on the midpoint estimation.

6.4.8 RSV group A and B co-detections

In 14 RSV infection episodes, one or more samples were identified to contain RSV group A and B. The timing of these co-detections was as shown in Figure 6.8. In most instances, (12/14) RSV group A was shed for longer duration relative to group B.

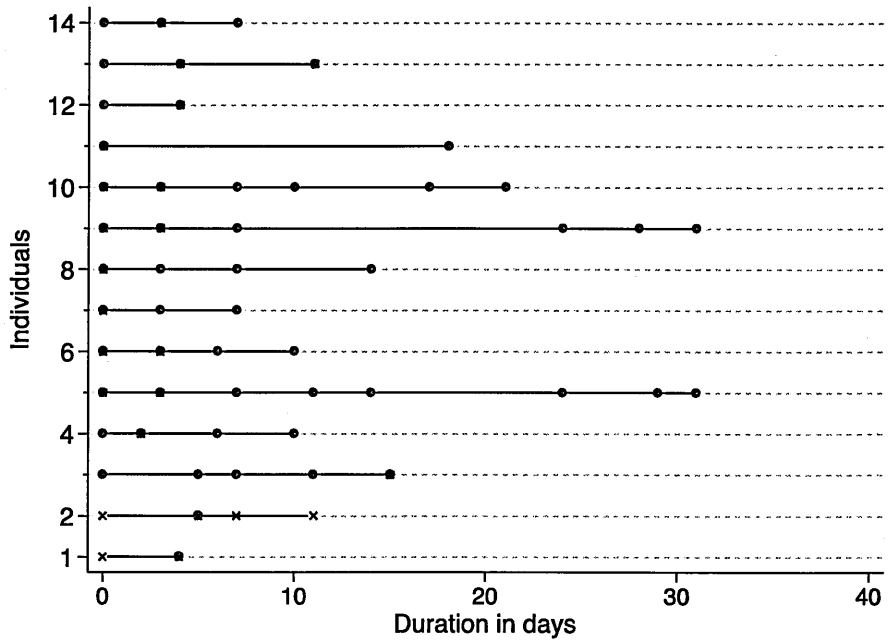


Figure 6.8: RSV shedding in episodes associated with co-detections of RSV group A and B. The filled red circle and blue marker (x) represents RSV group A and B, respectively. The red line links PCR positive samples denoting the RSV episode

6.4.9 Sequencing of the RSV G gene from individuals with 'suspected' reinfections with same RSV group

In 17 (9.5%) of the infected individuals, we observed an apparent 'suspected' repeat infection with same RSV group: 6 with RSV group A and 11 with RSV group B. The two or more phases of RSV shedding are shown in Figure 6.9. The mean age at time of the first infection for individuals with RSV group A and group B was 2.3 and 7.2 years, respectively. The duration between the positive samples ranged from 17 to 54 days with median of 28 days. Sequencing of the RSV G gene was successful in 10 (59%) of the 18 possible pairs of samples; one individual had three suspected RSV episodes (Table 6.4). The failure to sequence was mainly in samples with a Ct value of >28.0 by M-PCR, an indicator of low viral load, explaining the difficulty in

sequencing. Only one of the successfully sequenced paired samples had nucleotide differences: 13 nucleotide differences associated with three non-synonymous changes and a change in the stop codon position. The episodes in this individual (ID 1803 in Figure 6.9)) occurred 54 days apart.

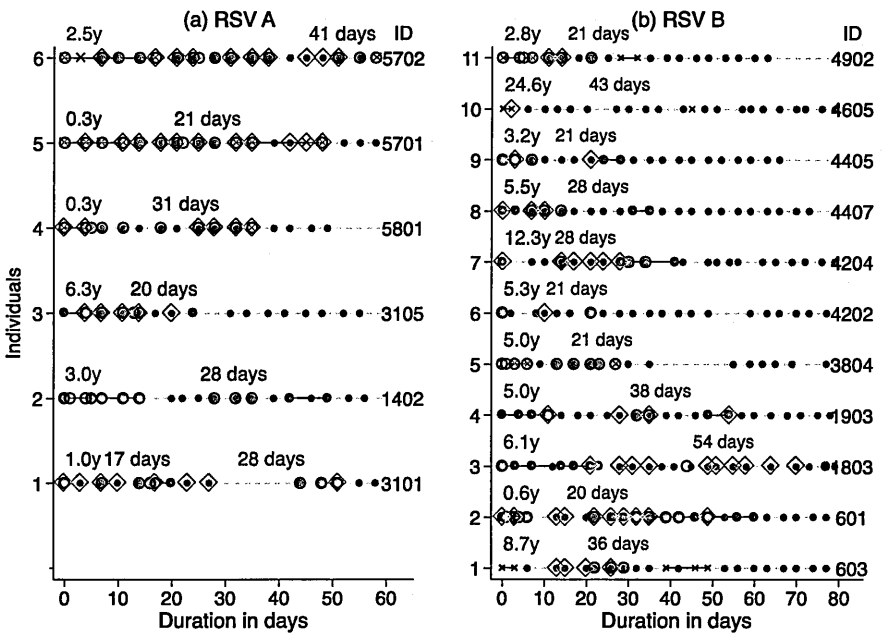


Figure 6.9: Infection episodes for individuals with suspected repeat infection with the same RSV group showing RSV G gene variability. The red crosses indicate RSV positives samples; filled red or blue circles show one or more RSV positives samples from each 'episode' were successfully sequenced. If red is in both episodes then the variants in each episode were identical and if one episode is red and the other blue then they were non-identical (Table 6.7). Green filled circles mean samples were collected but RSV negative. Pink diamonds represent detection of other viruses; while hollow black circles show when the individual had respiratory symptoms. The values in blue are ages at the start of the first episodes and the 'ID' represents the unique identification of the individuals. The intervals between the episodes are also shown in the graph region in days

Table 6.4: RSV G gene variability in 17-paired PCR positive samples from 17 individuals with suspected repeat infections

ID[1]	*RSV Group*	*Sequenced episode*			*G gene identical or different?*	*No. nt changes*
		First	Second	Third		
3101	A	No	Yes	Yes	Identical[2]	-
1402	A	Yes	Yes	-	Identical	-
3105	A	Yes	Yes	-	Identical	-
5801	A	Yes	No	-	-	-
5701	A	Yes	No	-	-	-
5702	A	Yes	No	-	-	-
0603	B	Yes	No	-	-	-
0601	B	Yes	Yes	-	Identical	-
1803	B	Yes	Yes	-	Different	12[3]
1903	B	Yes	Yes	-	Identical	-
3804	B	Yes	No	-	-	-
4202	B	Yes	Yes	-	Identical[2]	-
4204	B	No	No	-	-	-
4407	B	Yes	Yes	-	Identical	-
4405	B	Yes	Yes	-	Identical	-
4605	B	Yes	No	-		-
4902	B	Yes	Yes	-	Identical	-

Key: 1, the values in this column denote the individual identity with the first two digits denotes the household ID followed by the individual number; 2, the second episode sequence has a mixed population of C or T at one of the positions; 3, the differences between the two strains involves three non-synonymous changes and a change in the stop codon position

6.5 Discussion

These analyses involved 179 individuals from 40 households who had regular and frequent nasopharyngeal swab collections regardless of symptoms during a single RSV epidemic experiencing 205 infection episodes. We report a mean duration of RSV shedding of between 8.6 to 14.0 days based on the most conservative to least conservative approach. A more realistic estimate of 11.2 (95% CI, 10.1 – 12.3) days was based on the assumption that shedding starts midway between the date of the last negative sample and the first positive sample and ends midway between last positive sample and the next negative sample. The median duration of shedding using this approach was 8.5 (95% CI 5.5 -14.0) days with a range of 1 to 44 days. Okiro *et al* conducted a similar study in Kenya recruiting family members of infants involved in the Kilifi birth cohort (Okiro *et al.* 2010). Study cases were identified as RSV positive by Immunofluorescence antibody test (IFAT) on nasal washes collected from individuals with symptoms of acute respiratory infections and identified through weekly home visits (Okiro *et al.* 2010). These cases were then repeatedly sampled by nasal washing every 2-3 days until the first IFAT negative sample. A mean duration of shedding of 4.5 days was estimated from 193 infected children in 151 families (Okiro *et al.* 2010), compared with the corresponding estimate of 13.5 days from the current study. In a subset of the children whose start of symptoms could be established from the clinic records, the authors reported longer duration of 7.7 (6.4 – 9.0) days (Okiro *et al.* 2010). Given that RSV shedding has been reported before start of illness (Frank *et al.* 1981) the actual duration in the symptomatic children would have been an underestimate and our estimate of 13.5 days is likely to be more accurate. Other limitations of the Okiro's study were the lower sensitivity of IFAT compared to M-PCR used in the present study, and the terminations of sampling at the

first negative follow up sample. The present study revealed instances where negative samples arose within RSV episodes. Even though this observation raises questions on the relationship between infectivity and shedding duration, accounting for periods of RSV negative samples would still result to longer shedding duration compared to previous estimates. Alternative estimation of the shedding patterns by calculating the area under the Ct (viral load) curve would have some additional advantages and would be explored in future.

The Rochester family study with similar design as the present study (collecting samples every 3 to 4 days regardless of symptoms) reported lower estimates of duration of shedding of 3.4 to 7.4 days (Hall *et al.* 1976). However, the Hall *et al* study used the less sensitive screening method of viral culture, and most individuals were shedding the virus at the start of sampling. As a counter argument, it is not known to what degree PCR positivity equates with shedding of viable and infectious virus. Thus while molecular methods might be more sensitive, the resultant increase in duration of shedding over more traditional methods such as culture (which directly measures viral infectivity) may not necessarily translate to increased period of infectivity. Further work is warranted.

Of the previous studies noted, prolonged shedders for more than 3 weeks were reported (Hall *et al.* 1976; Frank *et al.* 1981). In the present study, 24/179 (13.4%) individuals shed RSV for more than 3 weeks whose median age at infection was 14.7 months. Most (22, 91.7%) of these episodes were symptomatic. It is not clear why the individuals had poor viral clearance and we did not test for HIV. Individuals with compromised immunity due to HIV infection or underlying medical problems (transplant receipts), have been known to shed RSV for longer (Madhi *et al.* 2000). In settings where HIV prevalence is high, the effect of the prolonged shedders might

influence the temporal epidemiology possibly continuing the spread of the virus during the low season (Madhi *et al.* 2000). Further studies are required to assess the epidemiological implications of the prolonged shedders.

More than two RSV episodes were experienced in 24 (13.4%) individuals in the single RSV epidemic. On average, the episodes were 4 weeks apart. Most of the suspected 'repeat' infections were with homologous RSV group (17/24. 70.8%). In the heterologous scenario, all except one were with RSV group B in the first episode followed by group A for the homologous infections, we sequenced the long ectodomain region of the G gene which encodes for the antigenically variable region of the attachment protein. Only 10/17 (59%) of the paired samples for the 'suspected' repeat infections with the same RSV group were successfully sequenced. Most (9/10) of the successfully sequenced infection-pairs had the same RSV variant. The second episodes occurred within 3 to 4 weeks interval. It is not apparent whether the two phases of RSV infection were clearly repeat infections or were persistence RSV infection with periods of low viral load undetectable from nasopharynx or by the M-PCR method. Using postmoterm lung tissue from infants, RSV RNA has been detected even in children dying during inter-epidemic periods suggesting persistence of RSV in lungs of these infants. Additionally, in an experimental RSV infection in mice, RSV RNA was found in lungs for more than 3 months after initial infection (Schwarze *et al.* 2004). Infective virus was detected albeit at low levels from lungs in mice subjected to T-cell depletion initiated over 4 months after the experimental infection. Similarly in experimental infection of rhesus macaques with measles virus, RNA was shown to persist in blood, respiratory tract or lymph node during intervening periods of non-infectiousness (Lin *et al.* 2012). The mechanism of RSV persistence is still unclear but there are suggestions it might be by means of low-grade

replication (Schwarze *et al.* 2004). The lack of variability in the virus identified in the two phases of infections suggests virus mutation might not be the primary mechanism for virus persistence or reinfection. The two phases of RSV shedding regardless of whether they are persistent infections or reinfection within the same RSV epidemic represents an interesting phenomenon of RSV which has potential importance on our existing view of acute RSV infection, the development of prolonged immunity and viral transmission and would benefit from further investigation.

Of the covariates examined, age, infection severity, presence of other viruses before and during RSV infection, and concurrent RSV infections in the household were significant determinants of the duration of shedding. Infants shed on average for 18 days indicating a low recovery rate relative to their older counterparts who shed for less than 12 days. The rate of recovery increased by age with individuals within 1 – 4 years, 5 – 15 years and ≥15 years age groups recovering 2.16, 3.40 and 2.56 times faster than the infants, respectively. The Rochester family study reported similar findings where longer shedding was observed for children <2 years compared to those aged 2 – 16 years (9 versus 4 days). The Okiro *et al* study did not find any association of age and duration of shedding but found that children with previous RSV infection (including those older than 3 years) had 37% increased recovery rate relative to those with no history of infection. In the present study, a subsequent RSV infection during the same RSV season was not significantly associated with reduced shedding duration. However, only 24 individuals experienced more than one RSV episode, a small sample size to allow detection of even large differences. The current study is of a shorter duration, so that distinguishing between primary and subsequent infections is not possible, so age must act as a proxy of previous exposure to RSV. Prolonged shedding enhances the possibility of person-person transmission and makes young

children a potential source of community spread of infection. These have important implications in the control and prevention of RSV infection.

Symptomatic RSV infections were associated with 43% reduced recovery rate relative to asymptomatic infections. A study involving 23 hospitalized children aged less than 2 years who were followed even after discharge reported an association of duration of shedding and infection severity (Hall *et al.* 1976). Children with lower respiratory tract infection shed for longer than those with upper respiratory tract infection (8.4 vs. 1.4 days). Duration of shedding may be related to severity of disease but evidence is controversial on the link between disease severity and viral load (Hall *et al.* 1975; Devincenzo 2004; Kuypers *et al.* 2004; DeVincenzo *et al.* 2005).

No published data exists on the effect of other viruses and concurrent household infections on the shedding durations of RSV. Coronavirus, rhinovirus and adenovirus were screened for all the nasal swabs collected. Presence of viral co-infection, defined by one or more viruses detected during the shedding period of RSV, was associated with 50% reduced recovery rate. The effect was similar even on examining the co-detected viruses i.e. coronavirus, rhinovirus and adenovirus, independently. In contrast, presence of other viruses before start of, but not during, the RSV episode was associated with an increased rate of cessation of RSV shedding by almost 50%, but with borderline significance. Individuals with co-infections might have had lower immunity hence the high susceptibility to multiple viral infections and low viral clearance relative to those experiencing only RSV infection. Conversely, individuals experiencing longer RSV infections would have greater chance of experiencing a concurrent viral infection. Recovery from a viral infection just prior to RSV might have led to up-regulation of innate viral immunity or non-specific cross-reactivity that reduced subsequent RSV shedding. While the possible mechanism of the observed

interference of prior viral infections on duration of RSV shedding is speculative, the interaction or interference of other viruses on RSV shedding patterns requires further assessment in a mathematical model framework to help tease out the direction of the interaction/effect and implications on virus transmission within households and the population.

Individuals from households where a third or more of the members had concurrent RSV infections had a reduced recovery rate by 46% than those from households with less than a third of the members infected. Close contacts in household might facilitate transmission of large viral inoculum resulting in more severe and prolonged shedding. It is impossible to rule out the role of super reinfections during the household RSV outbreaks.

The median RSV shedding durations presented in this study are generally longer than published estimates from non-hospital settings which could be attributed to methodological differences. The present study was informed by critical review of the previous studies and it incorporates frequent sampling regardless of symptoms and screening by highly sensitive PCR methods. Clearly, the use of the sensitive viral detection methods (PCR) results in longer estimates of shedding. However, more research is required to determine what level of shedding is related to infectiousness – it could be that highly sensitive methods are finding viral RNA, but that the level is too low for the virus to be effectively transmitted as recently observed in measles (Lin *et al.* 2012). Use of acceptable sample collection methods ensured wider involvement of all ages including older children and adults (Munywoki *et al.* 2011) and data on important covariates were systematically collected.

One of the main limitations of the present study arose from failure to determine the exact time of start and end of the shedding periods but analysis using the three assumptions helped to estimate the possible durations. Nevertheless, this did not affect the association of the explored factors with recovery rate (see Appendix Y). Short RSV infections occurring between specimen collections particularly in older individuals might have been missed but this is likely to be at random (non differential misclassification bias) leading to possible underestimation of the hazard ratios. Noteworthy is that midpoint imputation in the Cox model used in the current analysis would lead to the impression of greater precision in hazard ratio than is actually available from the data collected. Future work would be improved by directly taking the interval censoring into account in the models.

In conclusion, this study provides the best estimates of duration of RSV shedding and explores effects of various factors on recovery rate in naturally infected persons within a rural community setting. Of the covariates examined, age, infection severity, presence of other respiratory viruses and concurrent RSV infections in the household were significant factors related to the duration of shedding. The estimates of the RSV infectious durations are important in understanding the spread of infection in a population and will contribute to the development of transmission dynamic models to investigate the impact of interventions against RSV.

CHAPTER SEVEN

7 Estimation of Susceptibility and Transmission Parameters

7.1 Introduction

RSV is transmitted by large nasal droplets and fomites, with short survival time in the environment. This suggests that close contact is important for effective transmission. It follows that social interaction patterns (contacts and mixing patterns among individuals) is likely to have a direct impact on transmission dynamics of RSV. Households and schools are key social contexts with high contact rates providing environments with disproportionate importance for RSV transmission. Family studies indicate the importance of household size and school-age siblings as risk factors of RSV infection in infants (Hall *et al.* 1976; Okiro 2007). Various studies indicate RSV infection is frequently introduced into the home by siblings, resulting in high secondary attack rates within households (Berglund 1967; Hall *et al.* 1976; Okiro 2007), although a recent study reported contrasting findings (Crowcroft *et al.* 2008). Direct evidence of mother-to-child infection has been reported in one study from Guinea Bissau (Stensballe *et al.* 2004). It remains, however, that patterns of transmission, so important in determining the influence of infection (or control) in one age group on other groups, are poorly defined.

A prerequisite to assessing the impact of implementing various control strategies in prevention of RSV infections is the availability of realistic estimates of parameters that define the 'who acquires infection from whom' (WAIFW) matrix. Early work on childhood viruses concentrated on age-group related contact patterns (Schenzle 1984; Anderson and May 1991), and defined the WAIFW matrix by inference from age-stratified sero-prevalence or case notification data (Anderson and May 1985; Dietz

and Schenzle 1985; Grenfell and Anderson 1985). If there are n age groups, then the WAIFW matrix has n^2 elements, but there are only n estimates of the age-related rates of infection, so the WAIFW matrix is unidentifiable by this method without further assumptions. Three approaches have been proposed to overcome this problem. First, Medley suggested using multiple infections, each using the same WAIFW matrix (Farrington *et al.* 2001). Second, Edmunds *et al* pioneered the direct measurement of contact rates (Edmunds *et al.* 1997). Third, Kaanan *et al* used prior probabilities to inform the choice of matrix structure (Kanaan and Farrington 2005). It should be noted that heterogeneity in mixing patterns may not only be functions of age groupings, but also other forms of social stratification such as household structure and population demographics (e.g. urban or rural) – but these have not featured extensively in the development of models to predict vaccine impact (Anderson and May 1991). The use of contact data in transmission dynamics modelling requires the inference of the transmission probability following contact – which is a serious obstacle. Longitudinal studies can estimate this transmission probability more directly from more accurate, high-density observations of infections (Melegaro *et al.* 2004; Cooper *et al.* 2008). In the current study, we used the detailed RSV infection data from the household study, to directly estimate the relative susceptibility and transmission probabilities. Improved understanding of household transmission may inform on the merits of targeted control strategies.

7.2 Chapter outline

A transmission model incorporating time-dependency in community probability of infection, age-related susceptibility, household heterogeneity and contact structure is presented in this Chapter. Using the parameter estimates of transmission probability, relative susceptibility by age and post-infection, effects of household size and contact

structure in transmission, the relative contribution of various groups in the household in RSV acquisition are shown. The approach and results presented should be regarded as preliminary, and more development is required to reach publication standard.

7.3 Methods

7.3.1 Data

Data consist of the PCR detections of RSV in the household study. No distinction was made between RSV groups in the current analysis. Each household was considered independently. For the current analysis, we assumed the infection data were complete in the sense that all RSV episodes were observed, and the 'midpoint' approach described in Chapter 6 was used to identify the infection onsets and recovery times. Briefly, in the midpoint approach, infections were assumed to start (or end) on the date between the last negative and first positive (or last positive and first negative) samples if the interval between the two observations was 7 days or less. For instances where these intervals were greater than 7 days (i.e. left and right censored) a duration which was one half of the average interval for the other observations was subtracted from the first positive sampling date or added in the last positive sample date to estimate the start and end dates, respectively.

From the data, the shedding profile of each individual, i, was extracted, and denoted, $s_{iht} = 1$ if shedding on day t of the study, where h is the household of the individual, and $s_{iht} = 0$ if not shedding. The study had 493 individuals in 47 households and, in this analysis, were followed for 186 days from 1st December 2009 to 4th June 2010, i.e. $i = 1...493$, $t = 1...186$, and $h = 1...47$. Individuals are at risk of onset of shedding when not already shedding, i.e. $s_{iht} = 0$. The set of the days of onset of shedding for individual, i, are defined as U_{ih}. Not all individuals were within the study all the time. In particular, there are periods when individuals were absent: let $m_{ih}(t) = 1$ denote

when individuals were in the study, and $m_{ih}(t) = 0$ when they were absent or the active home follow-up had not yet begun. We assume that individuals who are absent cannot be infected by and cannot infect members of their household, and are at risk of infection due to community transmission only.

Table 7.1: The distribution of infection episodes between the 493 study individuals

Number of Episodes[1]	Number of Individuals
0	314
1	155
2	22
3	2

Key: 1, Infection episode defined as a period within which an individual provided specimens which were PCR positive for the same infecting RSV group with no more than 14 days separating any two positive samples (see Table 4.1)

There were 205 onsets of shedding in 91,598 person-days of observation with 2,200 (2.4%) days of shedding. In 24,380 (26.6%) days the participants were absent (i.e. away as defined in Table 3.1) and in 67,113 (73.3%) days the individuals were at risk of onset of shedding but did not. There were 5 individuals who were 'absent' in the 5 days preceding onset of shedding. There were 82 infection onsets which could only have been from the community i.e. there was no infection in the household in the 5 days preceding onset, or the individual was absent from the household prior to onset.

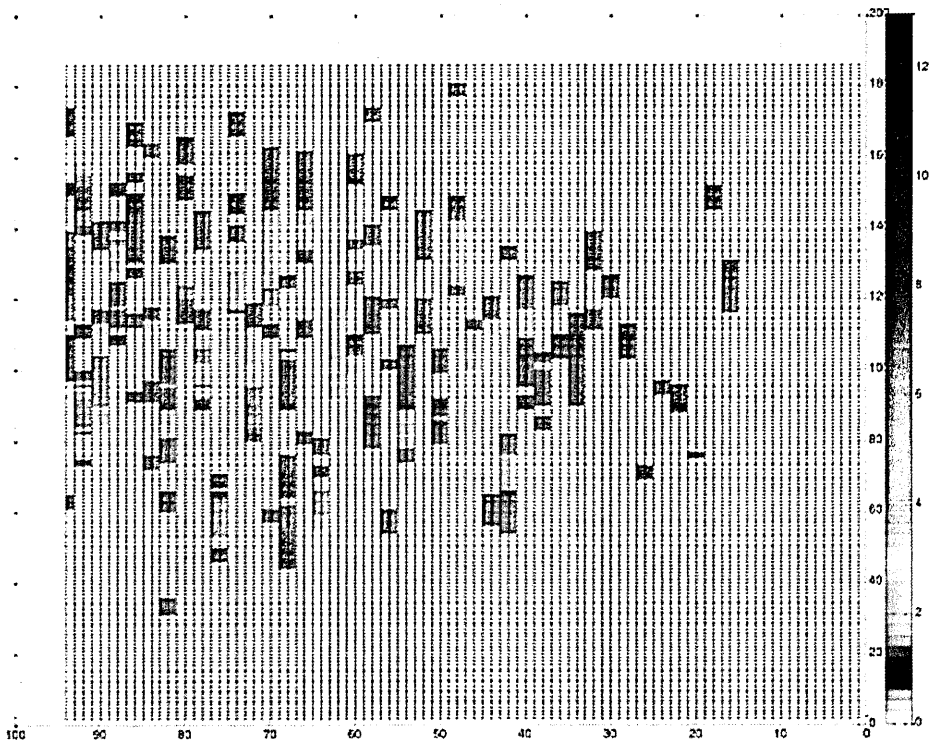


Figure 7.1: Household level pattern of infection. The most heavily infected household is on the left and time (in days) goes vertically. The colour scale shows the intensity of infections in each household and changes over time (in days)

Age at start of the study was taken as fixed covariates for each individual, denoted by A_i while previous occurrence of an observed RSV infection was set as a changing covariate over time which was denoted as B_{it}. For each pairing within the household, a set of covariates, which described the relationships or contact structure (e.g. mother-child, sibling-sibling etc.) were denoted as C_{ij}. For each household, the household size was taken as a covariate denoted as D_h. Covariates **A** and **B** define the state of an individual and were used to define relative susceptibility to the same virus exposure. Covariates **C** and **D** were used to define the relative transmission potential between groups of individuals. Covariates that define the infectiousness of individuals were

not included. The parameterization of the models was generally to make the study infants the reference group.

7.3.2 Basic model structure

For each individual on each day and for each virus, a variable that represents the individual's relative susceptibility to infection, π_{iht}, was constructed, which had the following structure:

$$\pi_{iht} = 0, if\ s_{iht} = 1$$
$$\pi_{iht} = exp(-\phi_A A_i - \gamma\phi_B B_{it}), if\ s_{iht} = 0 \quad \text{.....(i)}$$

Coefficients of the susceptibility covariates to be estimated were denoted by ϕ. This formulation does not admit super-infection during the shedding period. If there are no covariates being considered then relative susceptibility was 1 when an individual was not shedding, i.e. they were at full risk of infection when they were not shedding.

Let the rate of exposure for individual *i* in household *h* during day *t* be λ_{ih} (t):

$$\lambda_{ih}(t) = \sum_{j\neq i} \eta C_{ij}\ D_h s_{jht} + \epsilon(t) \text{.....(ii)}$$

where η are parameters relating to the covariates describing the relationships between individuals in the household and ε is the risk due to community exposure.

At a minimum level $C_{ij} = 1$, $D_h = 1$ and η determines the relationship between the amount of observed shedding and the rate of infection and:

$$\lambda_{ih}(t) = \eta \sum_{j\neq i}\ s_{jht} + \epsilon(t). \text{.....(iii)}$$

The contact structure in the household (e.g. infants and other groups) was considered, so that, in transmission matrix notation:

$$\eta C_{ij} = \eta_0 \begin{bmatrix} 1 & \eta_1 \\ \eta_1 & \eta_2 \end{bmatrix} \text{.....(iv)}$$

The probability that an individual became infected (i.e. enters the latent phase) during day t is then:

$$\alpha_{ih}(t) = \left(1 - e^{-\pi_{iht}\lambda_{ih}(t)}\right) \text{.....(v)}$$

This assumes a continuous relationship, i.e. there are no dose response effects, which would imply that higher doses/exposures would be more likely to lead to infection.

Because of the latent period (i.e. the time between infection and start of shedding), the probability of starting to shed on day t was taken as the weighted sum of the probability of infection in days prior to onset of shedding, where the weighting was the probability density function of the duration of the latent period:

$$p(i,h,t) = \sum_{j=0}^{j=m} \theta_j \alpha_{ih}(t-j) m_{ih}(t-j) \text{.....(vi)}$$

where the maximum latent period was m, and θ_j was the probability that the latent period was exactly j days (the θ have to sum to 1). Data is limited on RSV latency period following a natural infection and for this analysis we use data from an experimental challenge study in young healthy adults (Lee *et al.* 2004). In the challenge study, Lee *et al* demonstrated a latency period of 2–5 days after inoculation with 4, 4, 3 and 1 of the 12 infected adults starting to shed the virus at day 2, 3, 4, and 5 respectively. The θ were thus fixed at 0, 0.33, 0.33, 0.25 and 0.083 for day 1, 2, 3, 4, and 5 respectively.

7.4 Estimation

The model describes the infection process in terms of the observed exposure to virus (s), individual characteristics (constant, A, and varying, B), relatedness (C) and

household characteristics (constant D). The parameters determine individual susceptibility, both constant and varying (ϕ) and the relatedness between shedding and infection (η).

The total likelihood for an individual's observed data is:

$$L_i = \prod_{u \in U_{ih}} p(i,h,u) \prod_{u \notin U_{ih}} (1 - p(i,h,u)) \quad \text{.....(vii)}$$

where u denotes all the days of the study. This likelihood was maximized for each individual simultaneously. The model was developed and optimized in R Version 2.15.2 with the *optim* function (R Core Team 2012). The Broyden-Flethcher-Godfarb-Shanno (BFGS) method was used for numerical optimization (Broyden 1970; Fletcher 1970; Goldfarb 1970; Shanno 1970) after some trials. The R code used for estimation is included in the Appendix BB.

The time-dependent risk of community infection was estimated using MatLab curve fitting tool (http://www.mathworks.com/products/curvefitting/) (MATLAB, Release 2012), by fitting the sum of two Gaussian curves on the daily RSV detection data over the study period and normalising so that the maximum is unity.

The model was developed systematically starting with a simple model and adding in various components of parameter estimation. The parameters from the final model were taken forward in estimating the contribution of within household and community transmissions of RSV. In particular, the contributions of within household RSV acquisitions by relationships were estimated. For each observed onset, the force of infection for each household member and the community was weighted by the latency distribution and summed, and the contribution is expressed as a percentage. For

onsets when no household member was shedding during the latent period, the community has a 100% contribution.

Table 7.2: Summary of symbols and their definitions

Symbol	Definitions
i	Individual
h	Household
t	Time (days)
S	Shedding
U_{ih}	Set of onset days for each individual
$m_{ih}(t)$	Set of days an individual was present or absent from the household
A_i	Fixed covariates e.g. age at the start of the study
B_{it}	Time changing covariates e.g. occurrence of infection during follow up
C_{ij}	Contact structure in the household i.e. relationships like mother-infant
D_h	Household level covariates i.e household size
π_{iht}	Individual's relative susceptibility to infection
ϕ	Coefficients of the covariates
$\lambda_{ih}(t)$	The rate of exposure for individual i in household h during day t
η	η are parameters relating to the covariates describing the relationships between individuals in the household
ε	The risk due to community exposure
$\alpha_{ih}(t)$	Probability that an individual becomes infected during day t
β	Household infection rate
γ	a relative susceptibility post (observed) infection, γ

7.5 Results

7.5.1 Simplest model

The basic model had three parameters: constant community infection rate (ε), constant household infection rate (β) and a relative susceptibility post (observed) infection, γ. The results are given in Table 7.2.

7.5.2 Simple model with time-dependent community infection rate

The relative time-dependent risk of community-acquired infection is shown in Figure 7.2. The inclusion of this process improved the model fit (log-likelihood reduced from 1291.9 to 1265.9), and was included henceforth (Table 7.2). This modification, presumed that the study individuals were representative of the community pattern of shedding. Whilst the households were not a random sample of the community, the individuals might be more suitable for this purpose. The household individuals were not restricted in their interaction with the wider community and the prevalence of the detected RSV infections over time likely to be representative of the wave of RSV epidemic in the community.

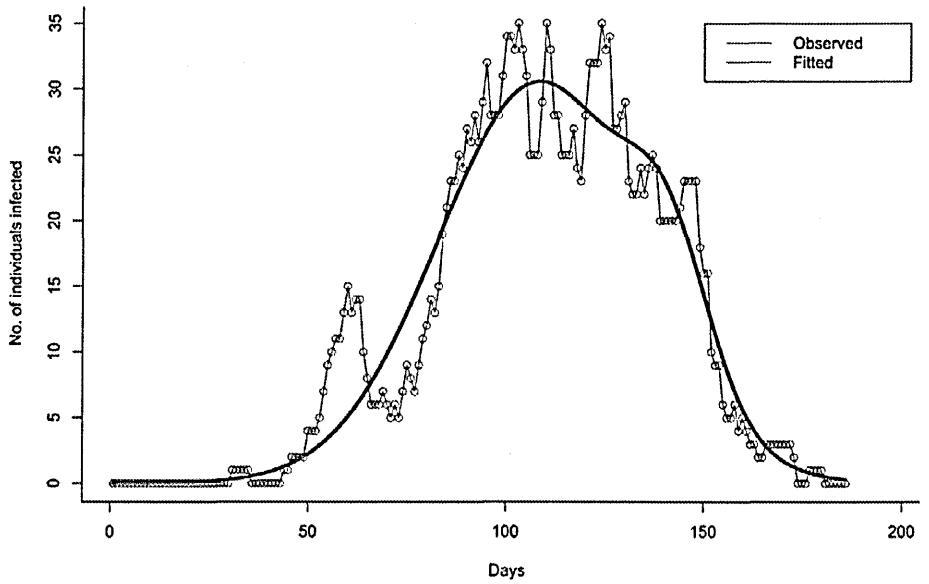


Figure 7.2: Density fit to time-dependent numbers of individuals shedding in the household study. Blue circles are the total number of individuals shedding each day. The red line is the fitted curve used in the model

7.5.3 Age-related susceptibility

Age is likely to change the risk of infection. The age (0 – 92 years) was divided into discrete groups and a single parameter fit to each (relative to the first age group which was considered as the <1year group, i.e. study infants and their young siblings). Inclusion of age-related susceptibility improved the model fit (Table 7.3). Using this model, the best fit was with five age groups. The relative risk of infection was 75.6%, 55.6%, 25.7% and 19.9 % in 1- 4, 5 – 14, 15 – 39 and ≥40 age groups compared to the infants (<1year). Figure 7.3 shows the relative susceptibility with age using the final model. The age susceptibility was considered when fitting the contact rate parameters.

Table 7.3: Model fitting results from basic and age-related susceptibility fitting

Model	# pars[1]	LL	ε	β	γ	Age[2]	Comments
Simple	3	1291.9	0.002	0.005	0.576	-	-
Simple[3]	3	1265.9	0.005	0.004	0.543	-	Time-dep.
2 age groups	4	1252.2	0.011	0.010	0.514	0.393	≥1 y
2 age groups	4	1242.0	0.009	0.009	0.460	0.381	≥5 y
2 age groups	4	1236.2	0.008	0.008	0.424	0.348	≥10 y
2 age groups	4	1243.8	0.006	0.006	0.464	0.377	≥15 y
3 age groups	5	1241.2	0.011	0.011	0.461	0.756	1 – <5 y
						0.321	≥5 y
4 age groups	6	1232.7	0.011	0.011	0.421	0.756	1 – <5 y
						0.556	5 – <15 y
						0.247	≥15 y
5 age groups	7	1232.4	0.011	0.011	0.422	0.756	1 – <5 y
						0.556	5 – <15 y
						0.257	15 – <40 y
						0.199	≥40 y

Key: 1, number of parameters estimated; 2, Age-related risk of infection, parameters estimates using the infants (<1 year olds) as the reference with 'complete' susceptibility; 3, simple model with time dependent community risk of infection; LL, log likelihood; ε, community risk of infection; β, within household probability of transmission; γ, post-infection susceptibility; y, age in years

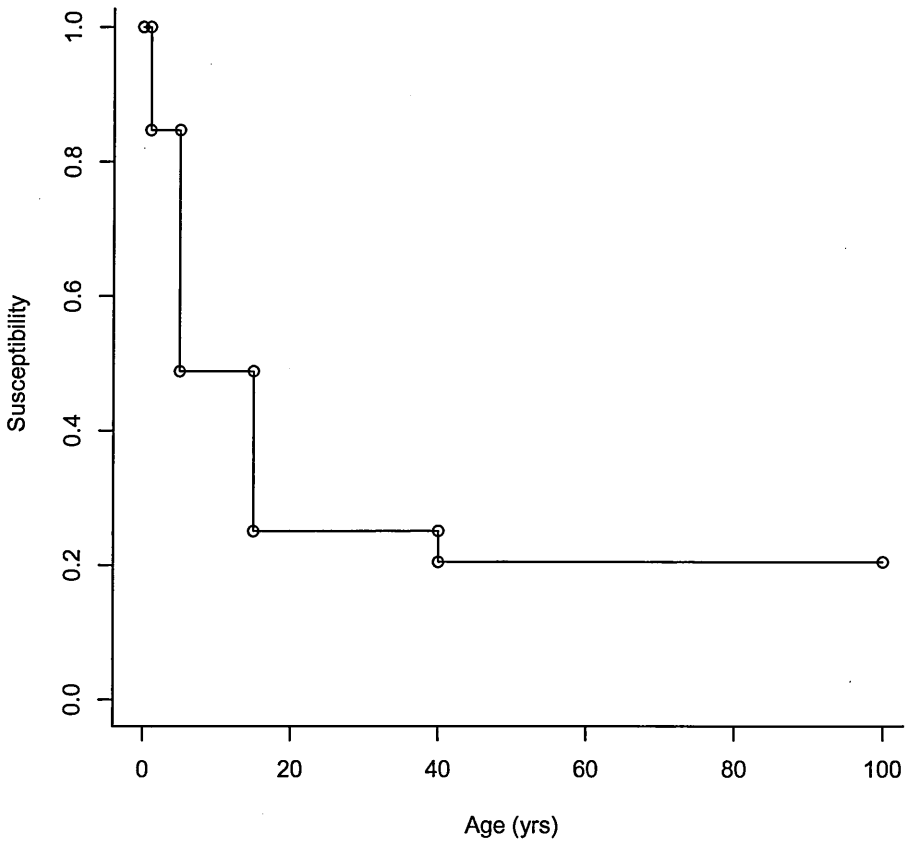


Figure 7.3: Age-related susceptibility parameters. Age was categorized into <1 year, 1-<5 years, 5-<15years and 15 <40years and ≥40 years. Based on estimates from the final model presented in Table 7.6

7.5.4 Household heterogeneity

Heterogeneity between households manifests itself as a difference in relative contact rate. This could be through many environmental factors, but we are only including size currently to look for potential density type effects. The households were grouped with the effect being relative to the smallest households. Multiple attempts using group definitions of household size suggested a single threshold effect with households of eight or more members having a lower relative transmission coefficient, Table 7.4. The effect was a reduction of transmission probability by 63.2%. The household size effect did not appear to interact with the age effect (Table 7.4).

Table 7.4. Results from household heterogeneity and age-related susceptibility fitting

Model	*# pars*[1]	*LL*	*ε*	*β*	*γ*	*h*	*Age*[2]	*Comments*
Base[3]	3	1265.9	0.005	0.004	0.543	-	-	
HH[4]	4	1260.2	0.005	0.011	0.513	0.321	-	HH≥8
5 age groups	7	1235.8	0.011	0.012	0.420	-	0.751 0.423 0.217 0.191	Age limits at 1, 5, 15, 40
HH & age	8	1231.1	0.011	0.028	0.406	0.368	0.746 0.434 0.224 0.187	HH & 5 age groups

Key: 1, number of parameters; 2, Age-related risk of infection, parameters estimates using the infants (<1 year olds) as the reference with 'complete' susceptibility; 3, simple model with time dependent community risk of infection; LL, log likelihood; 4, household (HH) effect included; ε, community risk of infection; β, within household probability of transmission; γ, post-infection susceptibility; h, transmission probability in households with 8 or more household members

7.5.5 Estimating transmission parameters

The contact structure was developed by categorizing household members based on the relationship to the study infant into infants, siblings, cousins, mother, father, and other adult members. In this Chapter, the infants also included any other child in the household below the age of one year at the start of the study or newborn during the study period. The numbers of different contacts (i.e. where individuals share a household) are given in the table below (Table 7.5). Thus, there are 27 pair-wise contacts between infants, and 296 pair-wise contacts between infants and siblings.

Table 7.5: Number of pair-wise contacts within households

	Infant	Sibling	Cousin	Mother	Father	Gparents	Other adults
Infant	27	296	118	99	63	34	68
Sibling		434	351	290	189	92	160
Cousin			238	129	64	52	193
Mother				37	67	42	70
Father					11	23	37
Gparents						6	44
Other adults							82

Key: Gparent, grand parents – this category was included in the 'other adults' in the subsequent analysis

The relationship categories are co-linear with age (see age distributions by relationships in Chapter 4), which was used to determine susceptibility, and were highly variable by household. For the purposes of fitting, relationships were grouped, e.g. nuclear family (study infant, sibling, mother, father) versus the other household members to assess whether nuclear families have different infection rates compared to non-parent-child relationships; children (infants, siblings and cousins) versus adults (mother, father, grandparents and other household members) to explore differences in transmission probability in adults relative to children. The possibilities for two group models were tried exhaustively with parameter estimates from the best fitting models shown in Table 7.6.

Table 7.6: Results from models exploring different contact structures

Model	# pars[1]	LL	ε	β x1000	γ	Comments
Base[2]	3	1265.9	0.005	4	0.543	Base
Infant/ others[3]	6	1256.5	0.005	3.0 17.9 8.3 2.8	0.500	2 groups
Child / adult[4]	6	1258.5	0.005	6.0 6.0 1.6 3.1	0.504	2 groups
Family / other[5]	6	1260.7	0.005	7.0 1.8 4.1 2.6	0.513	2 groups
Infant/child/adults[6]	11	1250.5	0.005	2.0 2.8 20.8 26.8 8.4 4.0 4.1 5.8 0.7	0.465	3 groups

Key: 1, number of parameters; 2, baseline model with time dependent community risk of infection; 3, Others include sibling, cousin, mother, father, grandparents and other adults in the household; 4, children included infants, sibling and cousins while adults had mother, father, grand parent and other adults in the household; 5, Family consisted of the nuclear family while other were the non-nuclear family members in the household (see text for details); 6, child included sibling and cousins while adults had mothers, fathers, grandparents and other adults in the household

7.5.6 Final models incorporating different contact structures

Taking the best age and household fits, the various contact groupings was incorporated to develop the final model (Table 7.7). There were small differences in the log-likelihood using different contact structures and the model with three contact groups (infants, children (siblings/cousins) and adults (mothers, fathers, grandparents and other adult members)) was taken forward.

Table 7.7: Final model fits

Model	*#pars*[1]	*LL*	*ε*	*β x1000*	*γ*	*Age*[2]	*h*
Infants / others	11	1228.0	0.010	0.38, 28.54	0.390	0.782	0.452
				32.28, 19.13		0.576	
						0.267	
						0.197	
Children / adults	11	1228.8	0.011	29.88, 23.96	0.391	0.735	0.377
				22.54, 19.42		0.541	
						0.267	
						0.206	
Family / other	11	1226.7	0.011	28.75, 8.50	0.389	0.730	0.545
				28.62, 14.79		0.524	
						0.245	
						0.176	
Infant, child &	16	1229.0	0.010	1041.9, 35.0, 28.2	0.396	*0.847*	0.514
Adults				25.2, 17.0, 16.2		*0.488*	
				52.6, 5.0, 37.4		*0.250*	
						0.205	

Key: 1, number of parameters

7.5.7 Sources of RSV infection in the study cohort

Of all the RSV episodes, 55.7% were attributable to community acquisitions (Table 7.8). Based on the model, 64.9%, 61.3%, 46.0%, 56.1%, 56.5% and 41.8% of the RSV infections among infants, siblings, cousins, mothers, fathers and other adults were from 'outside' the household. For each infection, the probability of infection from the groups in the household was calculated using the parameter estimates from

the final model with age and household effects and summed. Table 7.9 indicates the contribution to infection from each contact category within the household. When infants become infected, it was the older children particularly the siblings (63.0%) who were providing the majority of the risk (Figure 7.4). When siblings become infected, the risk was shared between the infants (35.7%), other siblings (44.2%) and cousins (10.0%). Similar pattern was observed for cousins, although mostly with each other (42.8%), and infants (32.2%) and siblings (12.9%). Mother acquired their infections from infants (71.8%) same as fathers (67.7%). Other adults in the household had a widely spread source of infections probably reflecting that they were a more heterogeneous grouping. Overall, infants, siblings, cousins, mother, father, and other adults contributed 34.9, 28.4, 21.4, 5.8, 1.4 and 8.2 of all RSV transmissions occurring within the household.

Table 7.8: Sources of RSV infection in the study cohort based on the model with three contact structures

i) Number of infections

	N	***NI***	***NIS***	***NIC***	***I***	***S***	***C***	***M***	***F***	***O***	***Com***
Infant	47	20	34	11	0.0	7.5	1.7	1.3	0.7	0.7	22.1
Sibling	164	47	79	20	10.9	13.5	3.1	2.7	0.0	0.4	48.4
Cousin	124	41	55	15	9.6	3.8	12.7	0.7	0.2	2.7	25.3
Mother	46	22	14	11	4.4	0.6	0.3	0.0	0.0	0.8	7.9
Father	33	11	6	8	1.8	0.1	0.2	0.4	0.0	0.3	3.4
Other	79	38	17	18	5.0	0.2	1.5	0.2	0.3	2.6	7.1
All	***493***	***179***	***205***	***83***	***31.7***	***25.8***	***19.4***	***5.3***	***1.3***	***7.4***	***114.1***

ii) Percentage contribution

	N	*NI*	*NIS*	*NIC*	*I*	*S*	*C*	*M*	*F*	*O*	*Com*
Infant	47	20	34	11	0.0	22.1	5.1	3.8	2.1	2.0	64.9
Sibling	164	47	79	20	13.8	17.1	3.9	3.4	0.0	0.5	61.3
Cousin	124	41	55	15	17.4	7.0	23.1	1.3	0.5	4.8	46.0
Mother	46	22	14	11	31.6	4.6	1.9	0.0	0.1	5.8	56.1
Father	33	11	6	8	29.4	1.0	2.5	6.0	0.0	4.5	56.5
Other	79	38	17	18	29.4	1.3	8.9	1.4	1.8	15.5	41.8
All	***493***	***179***	***205***	***83***	***15.4***	***12.6***	***9.5***	***2.6***	***0.6***	***3.6***	***55.7***

Key: N, number of individuals; NI, number ever infected; NIS, number of infections; NIC, number infected from community i.e. Primary and co-primary cases; I, infant; S, sibling; C, cousin; M, mother; F, father; O, other adults; Com, Community; 1, the columns (of the contact groups) shows the source while the rows are the recipients of the RSV infection

Table 7.9: Who infects whom in the within household transmissions

Relations[1]	**Infant**	**Sibling**	**Cousin**	**Mother**	**Father**	**Others**
Infant	0.0	63.0	14.6	10.8	5.9	5.6
Sibling	35.7	44.2	10.0	8.8	0.0	1.2
Cousin	32.2	12.9	42.8	2.3	0.8	9.0
Mother	71.8	10.5	4.3	0.0	0.3	13.1
Father	67.7	2.3	5.8	13.8	0.0	10.4
Others	50.6	2.2	15.2	2.4	3.1	26.5
All[2]	***34.9***	***28.4***	***21.4***	***5.8***	***1.4***	***8.2***

Key: 1, the columns shows the source while the rows are the recipients of the infection; 2, the column sums are the contributions to infection from each group

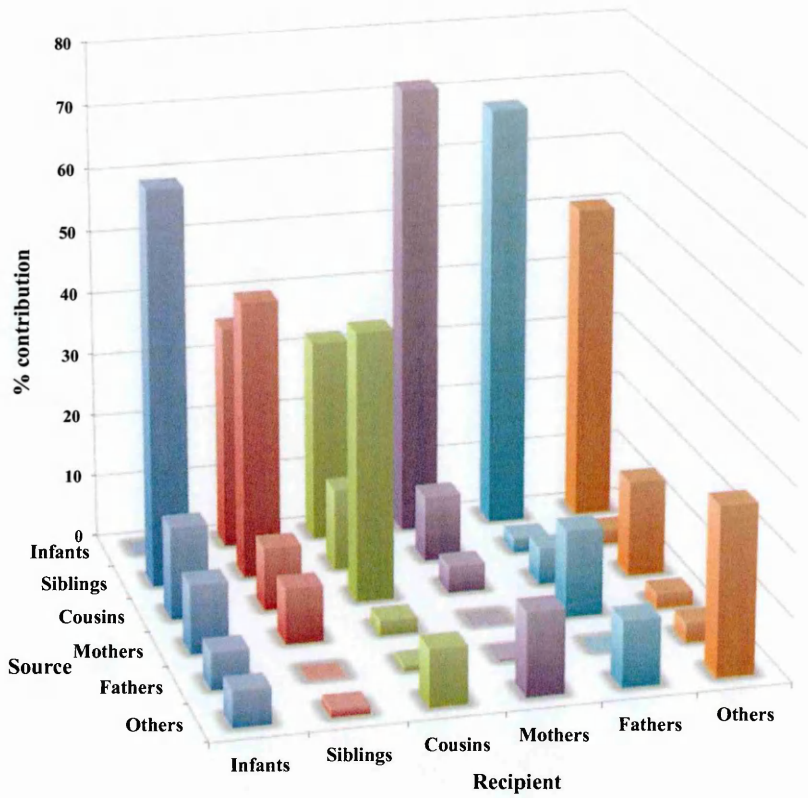


Figure 7.4: Who infects whom in within the household acquisition of RSV

7.6 Discussion

We report on a model aimed at estimating transmission and susceptibility parameters. Using the infants as the reference, we observed a decrease in susceptibility with age to less than a quarter by 15 years of age. Susceptibility post-infection reduced to 39.6% relative to pre-infection level during the same epidemic. Transmission probabilities were relatively higher in smaller households (with less than 8 members) compared to larger households. Similar findings have been reported for within-household transmission of influenza albeit with a lower threshold for classifying household size of less than 3 members (Cauchemez *et al.* 2009). Reanalysis of the Belgian POLYMOD data (Mossong *et al.* 2008) using a model for the contact network demonstrated small households (2-3 members) had higher contact rates relative to

larger households (Potter *et al.* 2011) which could partly explain the differential in transmission probability by household size. In addition, households with more than 8 members were more likely to have two or more nuclear families as described in Chapter 3.

On assessing the relative contribution of the various groups in the household on the observed RSV infections, our model indicated everybody had an equal likelihood of acquiring infection from the community (Table 7.8). This was a surprising finding, given that the age-related susceptibility changed so much. However, we believe that there might be an identifiability problem, in that the estimate of the level of community transmission necessarily includes some aspects of susceptibility (i.e. we can only observe infections that occur), and also includes aspects of community mixing (e.g. infections between school children). An alternative approach would be to just model transmission within households (i.e. ignoring the risk from the community). The current results, however, suggest that infants (who are most susceptible) don't mix as much with the remainder of the community. Siblings and cousins were the most common source of infant infections in within household acquisitions of RSV. Infected infants were a much greater risk for most of the household members especially mothers, fathers and other adults, contributing over 50% of within household infections in these categories. Overall, siblings and cousins (64.6%) were more likely to introduce infection into the households, which is in some part due to their numerical importance within households. Once RSV was in the household, children spread the infection amongst themselves, and the young children. Infants were infectious to the other household members but especially to their parents and other adults. This could be due to the prolonged duration of virus shedding (and high viral load) and close contacts associated with this age group.

The developed transmission model was based on a number of assumptions with varying limitations. A strong assumption was made in terms of the RSV infection data being 100% accurate, no dose effects in exposure and the invariant latency period. The actual data were interval-censored suggesting infections with short shedding durations (less than 4 days) were likely to be missed. Given that shedding duration was associated with age (see Chapter 5), there is a high propensity for the short infection episodes to be biased to the older children and adults. This is likely to result in failure to identify some introductions into the household and chains of transmission within the household. On timing of the infections, if two individuals started shedding on the same day and were the first to be infected in the household, the current framework identified both as infected from the community. Given the 3-4 day sampling interval and the short latency period, there is reasonable likelihood that one infected the other in reality, but this would be missed in the current model. If infants had a one-day shorter latent period, then the results are likely to be significantly different. The fact that individuals could only be infected from household members known to be shedding means that only few infections was passed from children to infants, since infants were usually infected quite early in the course of household outbreaks. Future work on this would require a Bayesian framework with modules to impute the unobserved first and last days of shedding as missing data (Cooper *et al.* 2008). There were very few households with more than one infant, which could have limited the reliability of the model fit. However, use of other 'contact' groups such as older children or adults as the reference did not change the interpretation of the results. RSV group data were not explicitly used in the current model hence we were not able to assess any group specific protection following infection, and relatedness between viruses assessed by molecular sequence will be used to inform these

estimates. In Chapter 5, we reported that the presence of other respiratory viruses, before and/or during an RSV episode influenced the duration of shedding. It would be of considerable interest to incorporate the effect of other viruses on RSV transmission to check for any competition or interactions in terms of relative susceptibility and transmission probability, and use these viruses to support estimation of contact parameters. The ultimate aim is to use these data to parameterise mathematical models for assessing the impact of various vaccination strategies.

Even though we captured temporal variations of RSV epidemic by weighting with the community probability of infection, we ignored spatial diversity by assuming an equal rate of transmission over the study location. Nevertheless this is likely to be of minimal significance given the study was constrained in a small geographical area likely to have similar exposure to the virus. Our approach also assumed the rate of exposure was proportional to the number of people shedding and did not include any information on the amount of shedding. Further work would benefit from incorporation of Ct values as a proxy for the viral load. Lastly, our model focused on relations in the household but RSV transmission could be dependent on other social patterns such as gender and employment status which could easily be included in the current framework.

In developing the model further, a simulation model for the study to assess the sensitivity of our sampling regime in detection of RSV infections would be required. The estimates would help in identifying the infections, which could have occurred but were not detected in the current framework.

In conclusion, this preliminary work suggests preventing the infection from entering the household requires vaccination of the children while prevention of within

household transmission would need vaccination of the infants. Further work would be required to assess the role of alternate transmission links such as mother-to-infant.

CHAPTER EIGHT

8 Overall Discussion

8.1 Chapter outline

This chapter aims to present the key findings and their implications as well as identify areas of future research. Limitations of the study and areas of improvements are highlighted.

8.2 Summary of main findings and their implications

Using the detailed infection data arising from the screening of 16924 NPS collected during a single RSV epidemic in the community, we demonstrate school-going children are responsible for introducing a large proportion of RSV into the households where there is an RSV naïve infant. Only one study in the past has looked at the spread of RSV within households in detail (Hall *et al.* 1976). Older children were identified to be introducers of RSV into the US families with young infants by Hall et al (Hall *et al.* 1976). Other less comparable studies report that younger children acquire infection from school-aged children within the household (Berglund 1967; Fox *et al.* 1975; Okiro 2007), indicating a significant contribution of within-school transmission to the overall transmission of RSV. Mothers have also been implicated as a source of infant infections in the household in one study (Stensballe *et al.* 2004). Assuming vaccination will lead to at least a short-term immunity to reinfection, immunisation of older siblings would reduce the rate at which infection is introduced into households and would potentially protect infants from infection during their first epidemic. Results from analysis of transmission patterns, albeit preliminary, identify infants as important in connecting the chains of transmission within the household and direct transmission to the adults in the household. It is epidemiologically

plausible that older children get their infection from schools where there is increased transmission due to intense contacts and introduce the infection to the household. In the household, they easily pass on the infection to infants who are highly susceptible to RSV. The infant in turn transmits the infection across the other household members including the mother due to the prolonged shedding and close contacts associated with this age group. This finding implies synchronised immunisation of older children and pregnant mothers (for passive transfer of protective immunity to the newborn) prior to the RSV season might offer good protection to the infants. The impacts of such a strategy require further assessment using mathematical models and in purposive field trials.

Our study reported an appreciable risk of infection across all age groups, underscoring the importance of reinfections in community transmission. The infection rates in older children (1 -14years) and adults (≥15 years) were particularly high at 43.3% and 21.5% respectively. Mothers had higher attack rates compared to fathers (26.1% versus 18.1%). Our data not only demonstrate that individuals get reinfected, but also that those with reinfections go on to reinfect others and generate chains of transmission. Thus we clearly show the importance of reinfections in the process of community transmission. Infection in older population can have important implications for transmission, infection, and disease in infants as possible reservoirs of infection. Nevertheless, an important aspect worthy of further consideration is the role of young infants in transmission dynamics. It is possible the infants drive the infection in the household due to their longer duration of shedding thus there are potential knock on effects of protecting infants.

The risk of clinical infection was dependent on age. The proportion of subclinical infections increased with age, in line with published literature suggesting risk of

disease is higher in primary infections than reinfections (Henderson *et al.* 1979; Glezen *et al.* 1986; Nokes *et al.* 2008), and that the risk of disease decreases with increase in age (Ohuma *et al.* 2012). These findings have implications for design of future studies in the community as they highlight the need for collecting samples regardless of symptoms. The amount of the virus circulating in the community would be underestimated if the contribution of asymptomatic infections is overlooked, a potentially important aspect in assessing the impact of RSV vaccines on viral circulation. However, assessment of the viability (infectiousness) of the viruses detected in subclinical infections and the role of subclinical infection in transmission is warranted. It will be important to assess whether Ct values (amount of viral shedding) in the RSV detections have any influence on transmission.

Transmission of infection is dependent on duration of infectiousness. We estimated the duration of shedding and factors that are associated with infection recovery. Of the covariates examined, age, infection severity, presence of other respiratory viruses and concurrent RSV infections in the household were significant factors related to the duration of shedding. Duration of viral shedding decreased with age and was shorter in subclinical infections relative to symptomatic infections. We did not assess the viral load profile of RSV shedding in order to identify possible peak periods for virus transmission (probably when viral load is high) and this forms part of intended future analyses. About 10% of the infected individuals shed RSV for more than 3 weeks. Most of these prolonged shedders were young with no clustering by household. Previously researchers showed immune-compromised individuals shed virus for prolonged periods (Hall *et al.* 1986). Our study did not test for HIV infection due to practical and ethical reasons.

We also report discontinuous shedding of RSV in another 10% of the cases: the duration between the shedding periods was more than two weeks raising questions of whether it was same infection or a repeat infection. Efforts to sequence the hyper-variable region of the G gene did not conclusively resolve this question and further work is planned using next generation (deep) sequencing and whole genome sequencing at Sanger Institute (http://www.sanger.ac.uk/research/projects/virusgenomics/). The presence of prolonged shedders and/or reinfections within short intervals has a bearing on our understanding of the biology of RSV infections particularly on mechanisms of viral persistence and immunological basis of reinfection. Even if our observations of intervening negative samples or periods was due to limitations in our PCR assay sensitivity, it still points to fluctuations in viral load over the course of the infection episode which also have an impact on viral spread. These observations were possible due to collection of samples irrespective of clinical status again underscoring the merits of this approach in community studies.

Human rhinovirus, human coronaviruses (OC43, NL63 and 229E strains) and adenoviruses co-circulated with RSV (Figure 8.1). Not surprising the incidence of co-infection of these viruses with RSV was not uncommon. Possible interaction of these viruses is likely to occur as they compete for the same ecological niche (respiratory epithelial cells). The effect of viral infection on nasopharyngeal carriage of bacteria is also being investigated using data from this study.

8.3 Study limitations and areas for improvement

Despite the frequent sampling regime for NPS, infections with short durations (<3 days) were likely to have been missed. The probability of missing these infections was higher in older age groups relative to infant because of the differential NPS

compliance by age and the associated short durations of shedding in adults. NPS compliance was higher in infants compared to other relatives and fathers and this could have an effect on estimating the importance of these groups in transmission. Unfortunately, efforts to supplement the infection data from the OF proved unsuccessful for the RSV-specific ELISA assay or insensitive by M-PCR testing. The antibody data would have also helped in assessing the influence of pre-infection antibody titres on infection and disease. There is still need for a less invasive but sensitive specimen collection method that is acceptable and can be practical for frequent sampling (ideally daily) across the ages in the community.

The potential importance of HIV in modifying RSV infection could not be established, as HIV status was not determined in the current study. The reported prolonged RSV shedding in some participants could be due to HIV infection. It would be useful in future to follow the prolonged shedders to see what happened to them as they can be easily tracked using the demographic surveillance data.

The study households were biased to those with older children, but it would be interesting to study transmissions in households without older children to see how the transmission patterns vary. We also followed the households for only one RSV season. Including multiple seasons would have provided additional information on the risk of reinfection and identify changes in within-household spread of RSV as a result of previous transmission.

8.4 Future research

As a natural extension of this work, our group has set up collaboration with Sanger Institute in the United Kingdom to develop the method for deep sequencing and whole genome sequencing of the RSV positive samples. The aim is to assess the

genetic diversity of the circulating viruses as well as to identify molecular markers to use in identifying transmission events (i.e. chains of transmission) within the household. The accruing data will also allow assessment of genetic relatedness of the virus detected during the same infection episode and between infected individuals of the same household over time, enabling analysis based on reconstruction of transmission events (Cooper *et al.* 2012). These data will be used to assess the importance of molecular evolution of the virus within individuals and the household.

We can also use Ct values as a proxy measure of viral load, e.g. to estimate the relative amount of virus shed over the shedding period. The aim would be to assess the relationship between viral shedding and the probability of transmission.

There were disparities in NPS collection rates with higher compliance reported in infants and mothers compared to older children and fathers. The compliance for oral fluid was high and acceptable across all the ages. The utility of the OF in supplementing infection data will be explored further using RNA UltraSense kit (Invitrogen), one-step quantitative RT-PCR System designed to detect viruses in low quantities.

In order to generate data with daily infections status, a Bayesian algorithm to impute missing data are required(Cooper *et al.* 2008). Data arising from these methods would be useful in the calculation of realistic transmission parameters such as duration of shedding, age-related susceptibility and transmission probability in the community and within the household. Mathematical models such as individual-based models of the transmission of RSV within communities incorporating the accruing household infection data will be developed. The aim would be to assess the potential impact of control measures such as vaccination and social distancing (Milne *et al.* 2008; Kelso

et al. 2009) on RSV transmission, infection and disease. Of relevance to this study, would be an assessment of household-based approaches in delivery of vaccines such as immunisation of older siblings or pregnant mothers.

Similar analysis for the other common respiratory viruses such as rhinoviruses, coronaviruses and adenoviruses as those presented in this thesis will be undertaken. Initial focus will be on coronaviruses, which have shown clear within-household spread (Figure 8.1). Our study suggests an effect of other respiratory viruses on the duration of RSV shedding alluding to possible interaction or competition of these viruses. These possible interactions will be explored in a mathematical model framework.

Older ages had high attack rates bringing infections into the household. We do not know from where they get their infection hence a need for broader studies that include schools or school-going children. These future studies would benefit from use of acceptable and non-invasive specimen collection methods.

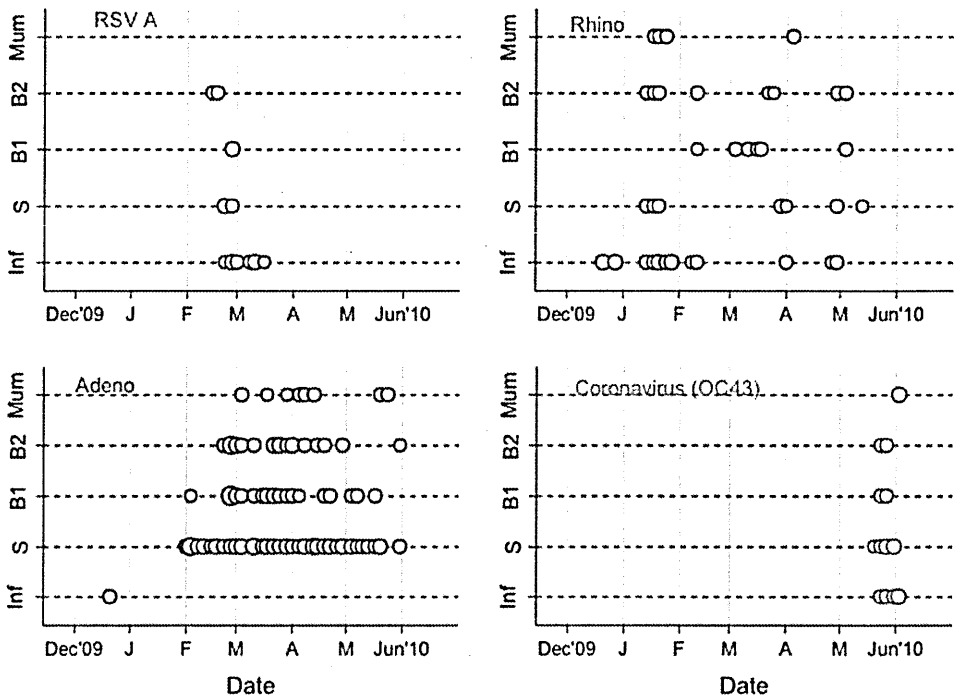


Figure 8.1: Infection patterns of a range of respiratory viruses in a household of 5 members, infant (Inf), one girl (S), two boys (B1 and B2) and their mother (Mum). In this household NPS were collected twice-a-week for the entire study period from all members. RSV A, respiratory syncytial virus group A; Rhino, Rhinovirus; Adeno, adenovirus; and human corona virus, OC43 strain

9 REFERENCES

Adegbola, R. A. (2012). "Childhood pneumonia as a global health priority and the strategic interest of the Bill & Melinda Gates Foundation." Clin Infect Dis **54 Suppl 2**: S89-92.

Agoti, C. N., A. G. Mwihuri, C. J. Sande, C. O. Onyango, G. F. Medley, P. A. Cane and D. J. Nokes (2012). "Genetic relatedness of infecting and reinfecting respiratory syncytial virus strains identified in a birth cohort from rural kenya." J Infect Dis **206**(10): 1532-1541.

Ahmed, J. A., M. A. Katz, E. Auko, M. K. Njenga, M. Weinberg, B. K. Kapella, *et al.* (2012). "Epidemiology of respiratory viral infections in two long-term refugee camps in Kenya, 2007-2010." BMC Infect Dis **12**(7): 7.

American Academy of Pediatrics (2012). Respiratory syncytial virus. . Elk Grove Villag, IL, American Academy of Pediatrics.

Anderson, L. J., P. R. Dormitzer, D. J. Nokes, R. Rappuoli, A. Roca and B. S. Graham (2013). "Strategic priorities for respiratory syncytial virus (RSV) vaccine development." Vaccine **31 Suppl 2**: B209-215.

Anderson, L. J., J. C. Hierholzer, C. Tsou, R. M. Hendry, B. F. Fernie, Y. Stone and K. McIntosh (1985). "Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies." J Infect Dis **151**(4): 626-633.

Anderson, R., Y. Huang and J. M. Langley (2010). "Prospects for defined epitope vaccines for respiratory syncytial virus." Future Microbiol **5**(4): 585-602.

Anderson, R. M. and B. T. Grenfell (1985). "Control of congenital rubella syndrome by mass vaccination." Lancet **2**(8459): 827-828.

Anderson, R. M. and B. T. Grenfell (1986). "Quantitative investigations of different vaccination policies for the control of congenital rubella syndrome (CRS) in the United Kingdom." J Hyg (Lond) **96**(2): 305-333.

Anderson, R. M. and R. M. May (1985). "Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes." Journal of Hygiene (Cambridge) **94**: 365-436.

Anderson, R. M. and R. M. May (1991). Infectious diseases of humans: dynamics and control. Oxford, Oxford University Press.

Badger, G. F., J. H. Dingle, A. E. Feller, R. G. Hodges, W. S. Jordan, Jr. and C. H. Rammelkamp, Jr. (1953). "A study of illness in a group of Cleveland families. III. Introduction of respiratory infections into families." Am J Hyg **58**(1): 41-46.

Baker, K. A. and M. E. Ryan (1999). "RSV infection in infants and young children. What's new in diagnosis, treatment, and prevention?" Postgrad Med **106**(7): 97-99, 103-104, 107-108 passim.

Becker, N. G. and K. Dietz (1995). "The effect of household distribution on transmission and control of highly infectious diseases." Mathematical Biosciences **127**(2): 207-219.

Bedoya, V. I., V. Abad and H. Trujillo (1996). "Frequency of respiratory syncytial virus in hospitalized infants with lower acute respiratory tract infection in Colombia." Pediatr Infect Dis J **15**(12): 1123-1124.

Berglund, B. (1967). "Respiratory syncytial virus infections in families. A study of family members of children hospitalized for acute respiratory disease." Acta Paediatr Scand **56**(4): 395-404.

Berkley, J. A., P. Munywoki, M. Ngama, S. Kazungu, J. Abwao, A. Bett, *et al.* (2010). "Viral etiology of severe pneumonia among Kenyan infants and children." Jama **303**(20): 2051-2057.

Berman, S. (1991). "Epidemiology of acute respiratory infections in children of developing countries." Rev Infect Dis **13 Suppl 6**: S454-462.

Berman, S., A. Duenas, A. Bedoya, V. Constain, S. Leon, I. Borrero and J. Murphy (1983). "Acute lower respiratory tract illnesses in Cali, Columbia: a two year ambulatory study." Pediatrics **71**: 210-218.

Bernstein, D. I., E. Malkin, N. Abughali, J. Falloon, T. Yi and F. Dubovsky (2012). "Phase 1 study of the safety and immunogenicity of a live, attenuated respiratory syncytial virus and parainfluenza virus type 3 vaccine in seronegative children." Pediatr Infect Dis J **31**(2): 109-114.

Borrero, I., L. Fajardo, A. Bedoya, A. Zea, F. Carmona and M. F. de Borrero (1990). "Acute respiratory tract infections among a birth cohort of children from Cali, Colombia, who were studied through 17 months of age." Rev Infect Dis **12 Suppl 8**: S950-956.

Botosso, V. F., P. M. Zanotto, M. Ueda, E. Arruda, A. E. Gilio, S. E. Vieira, *et al.* (2009). "Positive selection results in frequent reversible amino acid replacements in the G protein gene of human respiratory syncytial virus." PLoS Pathog **5**(1): e1000254.

Broyden, C. G. (1970). "The convergence of a class of double-rank minimization algorithms 1. general considerations." IMA Journal of Applied Mathematics **6**(1): 76-90.

Cane, P., Ed. (2007). Molecular Epidemiology and Evolution of Respiratory Syncytial Virus. Respiratory Syncytial Virus. 2007, Elsevier,.

Cane, P. and C. Pringle (1995). "Evolution of subgroup A respiratory syncytial virus: evidence for progressive accumulation of amino acid changes in the attachment protein." Journal of Virology **69**(5): 2918-2925.

Cane, P. and C. Pringle (1995). "Molecular epidemiology of human respiratory syncytial virus." Seminars in Virology **6**(6): 371-378.

Cane, P. A. (2001). "Molecular epidemiology of respiratory syncytial virus." Rev Med Virol **11**(2): 103-116.

Cane, P. A., D. A. Matthews and C. R. Pringle (1991). "Identification of variable domains of the attachment (G) protein of subgroup A respiratory syncytial viruses." Journal of General Virology **72**: 2091-2096.

Cane, P. A., D. A. Matthews and C. R. Pringle (1992). "Analysis of relatedness of subgroup A respiratory syncytial viruses isolated worldwide." Virus Res **25**(1-2): 15-22.

Cane, P. A., D. A. Matthews and C. R. Pringle (1994). "Analysis of respiratory syncytial virus strain variation in successive epidemics in one city." J Clin Microbiol **32**(1): 1-4.

Cane, P. A., M. Weber, M. Sanneh, R. Dackour, C. R. Pringle and H. Whittle (1999). "Molecular epidemiology of respiratory syncytial virus in The Gambia." Epidemiology and Infection **122**(1): 155-160.

Casiano-Colon, A. E., B. B. Hulbert, T. K. Mayer, E. E. Walsh and A. R. Falsey (2003). "Lack of sensitivity of rapid antigen tests for the diagnosis of respiratory syncytial virus infection in adults." J Clin Virol **28**(2): 169-174.

Cauchemez, S., C. A. Donnelly, C. Reed, A. C. Ghani, C. Fraser, C. K. Kent, *et al.* (2009). "Household transmission of 2009 pandemic influenza A (H1N1) virus in the United States." N Engl J Med **361**(27): 2619-2627.

Chan, K. H., J. S. Peiris, W. Lim, J. M. Nicholls and S. S. Chiu (2008). "Comparison of nasopharyngeal flocked swabs and aspirates for rapid diagnosis of respiratory viruses in children." J Clin Virol **42**(1): 65-69.

Chan, P. and A. Goh (1999). "Respiratory syncytial virus infection in young Malaysian children." Singapore Med J **40**(5): 336-340.

Chan, P. K., R. Y. Sung, K. S. Fung, M. Hui, K. W. Chik, F. A. Adeyemi-Doro and A. F. Cheng (1999). "Epidemiology of respiratory syncytial virus infection among

paediatric patients in Hong Kong: seasonality and disease impact." Epidemiol Infect **123**(2): 257-262.

Chanock, R., B. Roizman and R. Myers (1957). "Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). I. Isolation, properties and characterization." Am J Hyg **66**(3): 281-290.

Cherian, T., E. A. Simoes, M. C. Steinhoff, K. Chitra, M. John, P. Raghupathy and T. J. John (1990). "Bronchiolitis in tropical south India." Am J Dis Child **144**(9): 1026-1030.

Chew, F. T., S. Doraisingham, A. E. Ling, G. Kumarasinghe and B. W. Lee (1998). "Seasonal trends of viral respiratory tract infections in the tropics." Epidemiol Infect **121**(1): 121-128.

Coates, H. V., D. W. Alling and R. M. Chanock (1966). "An antigenic analysis of respiratory syncytial virus isolates by a plaque reduction neutralization test." Am J Epidemiol **83**(2): 299-313.

Collins, P. L., L. E. Dickens, A. Buckler-White, R. A. Olmsted, M. K. Spriggs, E. Camargo and K. V. Coelingh (1986). "Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order." Proc Natl Acad Sci U S A **83**(13): 4594-4598.

Collins, P. L. and B. R. Murphy, Eds. (2007). Vaccines against Human Respiratory Syncytial Virus. Respiratory Syncytial Virus. 2007, Elsevier,.

Connors, M., P. L. Collins, C. Y. Firestone and B. R. Murphy (1991). "Respiratory syncytial virus (RSV) F, G, M2 (22K), and N proteins each induce resistance to RSV challenge, but resistance induced by M2 and N proteins is relatively short-lived." J Virol **65**(3): 1634-1637.

Cooney, M. K., J. P. Fox and C. E. Hall (1975). "The Seattle Virus Watch. VI. Observations of infections with and illness due to parainfluenza, mumps and respiratory syncytial viruses and Mycoplasma pneumoniae." Am J Epidemiol **101**(6): 532-551.

Cooney, M. K., C. E. Hall and J. P. Fox (1972). "The Seattle virus watch. 3. Evaluation of isolation methods and summary of infections detected by virus isolations." Am J Epidemiol **96**(4): 286-305.

Cooper, B. S., T. Kypraios, R. Batra, D. Wyncoll, O. Tosas and J. D. Edgeworth (2012). "Quantifying type-specific reproduction numbers for nosocomial pathogens: evidence for heightened transmission of an Asian sequence type 239 MRSA clone." PLoS Comput Biol **8**(4): e1002454.

Cooper, B. S., G. F. Medley, S. J. Bradley and G. M. Scott (2008). "An augmented data method for the analysis of nosocomial infection data." Am J Epidemiol **168**(5): 548-557.

Crowcroft, N., F. Cutts and M. Zambon (1999). "Respiratory syncytial virus: an underestimated cause of respiratory infection, with prospects for a vaccine." Communicable Disease and Public Health **2**: 234-241.

Crowcroft, N. S., M. Zambon, T. G. Harrison, Q. Mok, P. Heath and E. Miller (2008). "Respiratory syncytial virus infection in infants admitted to paediatric intensive care units in London, and in their families." Eur J Pediatr **167**(4): 395-399.

Crowe, J. E., Jr. (2001). "Influence of maternal antibodies on neonatal immunization against respiratory viruses." Clin Infect Dis **33**(10): 1720-1727.

Cutts, F. T., S. M. Zaman, G. Enwere, S. Jaffar, O. S. Levine, J. B. Okoko, *et al.* (2005). "Efficacy of nine-valent pneumococcal conjugate vaccine against pneumonia and invasive pneumococcal disease in The Gambia: randomised, double-blind, placebo-controlled trial." Lancet **365**(9465): 1139-1146.

Dagan, R., D. Landau, H. Haikin and A. Tal (1993). "Hospitalization of Jewish and Bedouin infants in southern Israel for bronchiolitis caused by respiratory syncytial virus." Pediatr Infect Dis J **12**(5): 381-386.

Devincenzo, J. P. (2004). "Natural infection of infants with respiratory syncytial virus subgroups A and B: a study of frequency, disease severity, and viral load." Pediatr Res **56**(6): 914-917.

DeVincenzo, J. P. (2005). "Factors predicting childhood respiratory syncytial virus severity: what they indicate about pathogenesis." Pediatr Infect Dis J **24**(11 Suppl): S177-183, discussion S182.

DeVincenzo, J. P., J. Aitken and L. Harrison (2003). "Respiratory syncytial virus (RSV) loads in premature infants with and without prophylactic RSV fusion protein monoclonal antibody." J Pediatr **143**(1): 123-126.

DeVincenzo, J. P., C. M. El Saleeby and A. J. Bush (2005). "Respiratory syncytial virus load predicts disease severity in previously healthy infants." J Infect Dis **191**(11): 1861-1868.

Dietz, K. and D. Schenzle (1985). "Proportionate mixing models for age-dependent infection transmission." Journal of Mathematical Biology **22**(1): 117-120.

Djelantik, I. G. G., B. D. Gessner, S. Soewignjo, M. Steinhoff, A. Sutanto, A. Widjaya, *et al.* (2003). "Incidence and clinical features of hospitalization because of respiratory syncytial virus lower respiratory illness among children less than two years of age in a rural Asian setting." Pediatric Infectious Disease Journal **22**(2): 150-156.

Djelantik, I. G. G., B. D. Gessner, A. Sutanto, M. Steinhoff, M. Linehan, L. H. Moulton and S. Arjoso (2003). "Case fatality proportions and predictive factors for mortality among children hospitalized with severe pneumonia in a rural developing country setting." Journal of Tropical Pediatrics **49**(6): 327-332.

Durbin, A. P. and R. A. Karron (2003). "Progress in the development of respiratory syncytial virus and parainfluenza virus vaccines." Clin Infect Dis **37**(12): 1668-1677. Epub 2003 Nov 1620.

Edmunds, W. J., G. Kafatos, J. Wallinga and J. R. Mossong (2006). "Mixing patterns and the spread of close-contact infectious diseases." Emerg Themes Epidemiol **3**: 10.

Edmunds, W. J., C. J. O'Callaghan and D. J. Nokes (1997). "Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections." Proc R Soc Lond B **264**: 949-957.

Edmunds, W. J., C. J. O'Callaghan and D. J. Nokes (1997). "Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections." Proc Biol Sci **264**(1384): 949-957.

English, M. and J. A. Scott (2008). "What is the future for global case management guidelines for common childhood diseases?" PLoS Med **5**(12): e241.

Englund, J., W. P. Glezen and P. A. Piedra (1998). "Maternal immunization against viral disease." Vaccine **16**(14-15): 1456-1463.

Falsey, A. R., M. A. Formica, J. J. Treanor and E. E. Walsh (2003). "Comparison of quantitative reverse transcription-PCR to viral culture for assessment of respiratory syncytial virus shedding." J Clin Microbiol **41**(9): 4160-4165.

Falsey, A. R. and E. E. Walsh (2000). "Respiratory syncytial virus infection in adults." Clin Microbiol Rev **13**(3): 371-384.

Falsey, A. R., E. E. Walsh, J. Capellan, S. Gravenstein, M. Zambon, E. Yau, *et al.* (2008). "Comparison of the safety and immunogenicity of 2 respiratory syncytial virus (rsv) vaccines--nonadjuvanted vaccine or vaccine adjuvanted with alum--given concomitantly with influenza vaccine to high-risk elderly individuals." J Infect Dis **198**(9): 1317-1326.

Farrington, C. P., M. N. Kanaan and N. J. Gay (2001). "Estimation of the basic reproduction number for infectious diseases from age-stratified serological survey data." Journal of the Royal Statistical Society Series C-Applied Statistics **50**: 251-283.

Feikin, D. R., M. K. Njenga, G. Bigogo, B. Aura, G. Aol, A. Audi, *et al.* (2012). "Etiology and Incidence of viral and bacterial acute respiratory illness among older children and adults in rural western Kenya, 2007-2010." PLoS One **7**(8): e43656.

Felton, K., I. Pandya-Smith, A. Curns, A. Fry, L. Anderson and N. Keeler (2004). "Respiratory syncytial virus activity--United States, 2003-2004." MMWR Morb Mortal Wkly Rep **53**(49): 1159-1160.

Fishaut, M., D. Tubergen and K. McIntosh (1980). "Cellular response to respiratory viruses with particular reference to children with disorders of cell-mediated immunity." J Pediatr **96**(2): 179-186.

Fletcher, R. (1970). "A new approach to variable metric algorithms." The computer journal **13**(3): 317-322.

Forgie, I. M., K. P. O'Neill, N. Lloyd-Evans, M. Leinonen, H. Campbell, H. C. Whittle and B. M. Greenwood (1991). "Etiology of acute lower respiratory tract infections in Gambian children: II. Acute lower respiratory tract infection in children ages one to nine years presenting at the hospital." Pediatr Infect Dis J **10**(1): 42-47.

Fox, J. P., M. K. Cooney and C. E. Hall (1975). "The Seattle virus watch. V. Epidemiologic observations of rhinovirus infections, 1965-1969, in families with young children." Am J Epidemiol **101**(2): 122-143.

Fox, J. P., L. Elveback, W. Scott, L. Gatewood and E. Ackerman (1971). "Herd immunity: basic concept and relevance to public health immunization practices." Am J Epidemiol **94**(3): 179-189.

Fox, J. P. and C. E. Hall (1971). "Viruses in families." Lancet **1**(7711): 1240-1241.

Frank, A. L., L. H. Taber, C. R. Wells, J. M. Wells, W. P. Glezen and A. Paredes (1981). "Patterns of shedding of myxoviruses and paramyxoviruses in children." J Infect Dis **144**(5): 433-441.

Gilani, Z., Y. D. Kwong, O. S. Levine, M. Deloria-Knoll, J. A. Scott, K. L. O'Brien and D. R. Feikin (2012). "A literature review and survey of childhood pneumonia etiology studies: 2000-2010." Clin Infect Dis **54 Suppl 2**: S102-108.

Glenn, G. M., G. Smith, L. Fries, R. Raghunandan, H. Lu, B. Zhou, *et al.* (2013). "Safety and immunogenicity of a Sf9 insect cell-derived respiratory syncytial virus fusion protein nanoparticle vaccine." Vaccine **31**(3): 524-532.

Glezen, W. P., L. H. Taber, A. L. Frank and J. A. Kasel (1986). "Risk of primary infection and reinfection with respiratory syncytial virus." American Journal of Diseases of Children **140**: 543-546.

Goldfarb, D. (1970). "A family of variable metric methods derived by variational means." Mathematics of computation **24**(109): 23-26.

Goldstein, E., K. Paur, C. Fraser, E. Kenah, J. Wallinga and M. Lipsitch (2009). "Reproductive numbers, epidemic spread and control in a community of households." Math Biosci **221**(1): 11-25.

Gomez, M., M. A. Mufson, F. Dubovsky, C. Knightly, W. Zeng and G. Losonsky (2009). "Phase-I study MEDI-534, of a live, attenuated intranasal vaccine against respiratory syncytial virus and parainfluenza-3 virus in seropositive children." Pediatr Infect Dis J **28**(7): 655-658.

Gonzalez, I. M., R. A. Karron, M. Eichelberger, E. E. Walsh, V. W. Delagarza, R. Bennett, *et al.* (2000). "Evaluation of the live attenuated cpts 248/404 RSV vaccine in combination with a subunit RSV vaccine (PFP-2) in healthy young and older adults." Vaccine **18**(17): 1763-1772.

Grenfell, B. T. and R. M. Anderson (1985). "The estimation of age-related rates of infection from case notifications and serological data." Journal of Hygiene **95**: 419-436.

Guerguerian, A. M., M. Gauthier, M. H. Lebel, C. A. Farrell and J. Lacroix (1999). "Ribavirin in ventilated respiratory syncytial virus bronchiolitis. A randomized, placebo-controlled trial." Am J Respir Crit Care Med **160**(3): 829-834.

Gunson, R. N., T. C. Collins and W. F. Carman (2005). "Real-time RT-PCR detection of 12 respiratory viral infections in four triplex reactions." J Clin Virol **33**(4): 341-344.

Hall, C., E. Walsh, C. Long and K. Schnabel (1991). "Immunity to and frequency of reinfection with respiratory syncytial virus." Journal of Infectious Diseases **163**(4): 693-698.

Hall, C. B. (1977). "The shedding and spreading of respiratory syncytial virus." Pediatr Res **11**(3 Pt 2): 236-239.

Hall, C. B. and R. G. Douglas, Jr. (1981). "Modes of transmission of respiratory syncytial virus." J Pediatr **99**(1): 100-103.

Hall, C. B., R. G. Douglas, Jr. and J. M. Geiman (1975). "Quantitative shedding patterns of respiratory syncytial virus in infants." J Infect Dis **132**(2): 151-156.

Hall, C. B., R. G. Douglas, Jr. and J. M. Geiman (1976). "Respiratory syncytial virus infections in infants: quantitation and duration of shedding." J Pediatr **89**(1): 11-15.

Hall, C. B., R. G. Douglas, Jr. and J. M. Geiman (1980). "Possible transmission by fomites of respiratory syncytial virus." J Infect Dis **141**(1): 98-102.

Hall, C. B., R. G. Douglas, Jr., J. M. Geiman and M. K. Messner (1975). "Nosocomial respiratory syncytial virus infections." N Engl J Med **293**(26): 1343-1346.

Hall, C. B., R. G. Douglas, Jr., K. C. Schnabel and J. M. Geiman (1981). "Infectivity of respiratory syncytial virus by various routes of inoculation." Infect Immun **33**(3): 779-783.

Hall, C. B., J. M. Geiman, R. Biggar, D. I. Kotok, P. M. Hogan and G. R. Douglas, Jr. (1976). "Respiratory syncytial virus infections within families." N Engl J Med **294**(8): 414-419.

Hall, C. B., J. M. Geiman, R. G. Douglas, Jr. and M. P. Meagher (1978). "Control of nosocomial respiratory syncytial viral infections." Pediatrics **62**(5): 728-732.

Hall, C. B., C. E. Long and K. C. Schnabel (2001). "Respiratory syncytial virus infections in previously healthy working adults." Clin Infect Dis **33**(6): 792-796.

Hall, C. B., J. T. McBride, C. L. Gala, S. W. Hildreth and K. C. Schnabel (1985). "Ribavirin treatment of respiratory syncytial viral infection in infants with underlying cardiopulmonary disease." JAMA **254**(21): 3047-3051.

Hall, C. B., K. R. Powell, N. E. MacDonald, C. L. Gala, M. E. Menegus, S. C. Suffin and H. J. Cohen (1986). "Respiratory syncytial viral infection in children with compromised immune function." N Engl J Med **315**(2): 77-81.

Hall, C. B., E. E. Walsh, J. F. Hruska, R. F. Betts and W. J. Hall (1983). "Ribavirin treatment of experimental respiratory syncytial viral infection. A controlled double-blind study in young adults." JAMA **249**(19): 2666-2670.

Hall, C. B., G. A. Weinberg, M. K. Iwane, A. K. Blumkin, K. M. Edwards, M. A. Staat, *et al.* (2009). "The burden of respiratory syncytial virus infection in young children." N Engl J Med **360**(6): 588-598.

Hall, R. and N. G. Becker (1996). "Preventing epidemics in a community of households." Epidemiol Infect **117**(3): 443-455.

Hammitt, L. L., S. Kazungu, S. C. Morpeth, D. G. Gibson, B. Mvera, A. J. Brent, *et al.* (2012). "A preliminary study of pneumonia etiology among hospitalized children in Kenya." Clin Infect Dis **54 Suppl 2**(2): S190-199.

Hammitt, L. L., S. Kazungu, S. Welch, A. Bett, C. O. Onyango, R. N. Gunson, *et al.* (2011). "Added value of an oropharyngeal swab in detection of viruses in children hospitalized with lower respiratory tract infection." J Clin Microbiol **49**(6): 2318-2320.

Hemming, V. G., W. Rodriguez, H. W. Kim, C. D. Brandt, R. H. Parrott, B. Burch, *et al.* (1987). "Intravenous immunoglobulin treatment of respiratory syncytial virus infections in infants and young children." Antimicrob Agents Chemother **31**(12): 1882-1886.

Henderson, F., A. Collier, W. J. Clyde and F. Denny (1979). "Respiratory-syncytial-virus infections, reinfections and immunity. A prospective, longitudinal study in young children." New England Journal of Medicine **300**: 530-534.

Hendry, R. M., J. C. Burns, E. E. Walsh, B. S. Graham, P. F. Wright, V. G. Hemming, *et al.* (1988). "Strain-specific serum antibody responses in infants undergoing primary infection with respiratory syncytial virus." J Infect Dis **157**(4): 640-647.

House, T. and M. J. Keeling (2008). "Deterministic epidemic models with explicit household structure." Math Biosci **213**(1): 29-39.

House, T. and M. J. Keeling (2009). "Household structure and infectious disease transmission." Epidemiol Infect **137**(5): 654-661.

Hsu, K. H., M. D. Lubeck, A. R. Davis, R. A. Bhat, B. H. Selling, B. M. Bhat, *et al.* (1992). "Immunogenicity of recombinant adenovirus-respiratory syncytial virus vaccines with adenovirus types 4, 5, and 7 vectors in dogs and a chimpanzee." J Infect Dis **166**(4): 769-775.

Hussey, G. D., P. Apolles, Z. Arendse, J. Yeates, A. Robertson, G. Swingler and H. J. Zar (2000). "Respiratory syncytial virus infection in children hospitalised with acute lower respiratory tract infection." S Afr Med J **90**(5): 509-512.

Jin, H., X. Cheng, V. L. Traina-Dorge, H. J. Park, H. Zhou, K. Soike and G. Kemble (2003). "Evaluation of recombinant respiratory syncytial virus gene deletion mutants in African green monkeys for their potential as live attenuated vaccine candidates." Vaccine **21**(25-26): 3647-3652.

Johnson, P. R. and P. L. Collins (1988). "The fusion glycoproteins of human respiratory syncytial virus of subgroups A and B: sequence conservation provides a structural basis for antigenic relatedness." J Gen Virol **69 (Pt 10)**: 2623-2628.

Johnson, P. R., M. K. Spriggs, R. A. Olmsted and P. L. Collins (1987). "The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins." Proc Natl Acad Sci U S A **84**(16): 5625-5629.

Kaaijk, P., W. Luytjes and N. Y. Rots (2013). "Vaccination against RSV: Is maternal vaccination a good alternative to other approaches?" Hum Vaccin Immunother **9**(6).

Kanaan, M. N. and C. P. Farrington (2005). "Matrix models for childhood infections: a Bayesian approach with applications to rubella and mumps." Epidemiol Infect **133**(6): 1009-1021.

Karron, R. A., R. J. Singleton, L. Bulkow, A. Parkinson, D. Kruse, I. DeSmet, *et al.* (1999). "Severe respiratory syncytial virus disease in Alaska native children. RSV Alaska Study Group." J Infect Dis **180**(1): 41-49.

Karron, R. A., P. F. Wright, R. B. Belshe, B. Thumar, R. Casey, F. Newman, *et al.* (2005). "Identification of a recombinant live attenuated respiratory syncytial virus vaccine candidate that is highly attenuated in infants." J Infect Dis **191**(7): 1093-1104.

Karstaedt, A. S., M. Hopley, M. Wong, H. H. Crewe-Brown and A. Tasset-Tisseau (2009). "Influenza- and respiratory syncytial virus-associated adult mortality in Soweto." S Afr Med J **99**(10): 750-754.

Kelso, J. K., G. J. Milne and H. Kelly (2009). "Simulation suggests that rapid activation of social distancing can arrest epidemic development due to a novel strain of influenza." BMC Public Health **9**: 117.

Kenya National Bureau of Statistics (2010). 2009 Population and Housing Census Results. N. d. a. v. ministry of State for planning. http://www.knbs.or.ke/Census Results/Presentation by Minister for Planning revised.pdf, KNBS.

Kenya National Bureau of Statistics, I. M. (2010). Kenya Demographic and Health Survey 2008-09. Calverton, Maryland, KNBS and ICF Macro.

Kirkwood, B. R. and J. A. C. Sterne (2003). Essential Medical Statistics. Oxford, UK, Blackwell publishing.

Kravetz, H., V. Knight, R. Chanock, J. Morris, K. Johnson, D. Rifkind and J. Utz (1961). "Respiratory syncytial virus. III. Production of illness and clinical observations in adult volunteers." Journal of the American Medical Association, **176**: 657-663.

Kusel, M. M., N. H. de Klerk, P. G. Holt, T. Kebadze, S. L. Johnston and P. D. Sly (2006). "Role of respiratory viruses in acute upper and lower respiratory tract illness in the first year of life: a birth cohort study." Pediatr Infect Dis J **25**(8): 680-686.

Kuypers, J., N. Wright and R. Morrow (2004). "Evaluation of quantitative and type-specific real-time RT-PCR assays for detection of respiratory syncytial virus in respiratory specimens from children." J Clin Virol **31**(2): 123-129.

Langley, J. M., V. Sales, A. McGeer, R. Guasparini, G. Predy, W. Meekison, *et al.* (2009). "A dose-ranging study of a subunit Respiratory Syncytial Virus subtype A vaccine with and without aluminum phosphate adjuvantation in adults > or =65 years of age." Vaccine **27**(42): 5913-5919.

Law, B. J., E. E. Wang, N. MacDonald, J. McDonald, S. Dobson, F. Boucher, *et al.* (1997). "Does ribavirin impact on the hospital course of children with respiratory syncytial

virus (RSV) infection? An analysis using the pediatric investigators collaborative network on infections in Canada (PICNIC) RSV database." Pediatrics **99**(3): E7.

Lee, F. E., E. E. Walsh, A. R. Falsey, R. F. Betts and J. J. Treanor (2004). "Experimental infection of humans with A2 respiratory syncytial virus." Antiviral Res **63**(3): 191-196.

Levine, O. S., K. L. O'Brien, M. Deloria-Knoll, D. R. Murdoch, D. R. Feikin, A. N. DeLuca, *et al.* (2012). "The Pneumonia Etiology Research for Child Health Project: a 21st century childhood pneumonia etiology study." Clin Infect Dis **54 Suppl 2**: S93-101.

Lin, W. H., R. D. Kouyos, R. J. Adams, B. T. Grenfell and D. E. Griffin (2012). "Prolonged persistence of measles virus RNA is characteristic of primary infection dynamics." Proc Natl Acad Sci U S A **109**(37): 14989-14994.

Liu, L., H. L. Johnson, S. Cousens, J. Perin, S. Scott, J. E. Lawn, *et al.* (2012). "Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000." Lancet **379**(9832): 2151-2161.

Longini, I. M., Jr., J. S. Koopman, A. S. Monto and J. P. Fox (1982). "Estimating household and community transmission parameters for influenza." Am J Epidemiol **115**(5): 736-751.

Lozano, R., M. Naghavi, K. Foreman, S. Lim, K. Shibuya, V. Aboyans, *et al.* (2012). "Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010." Lancet **380**(9859): 2095-2128.

Machata, A. M., B. Kabon, H. Willschke, D. Prayer and P. Marhofer (2010). "Upper airway size and configuration during propofol-based sedation for magnetic resonance imaging: an analysis of 138 infants and children." Paediatr Anaesth **20**(11): 994-1000.

Madhi, S. A., A. Madhi, K. Petersen, M. Khoosal and K. P. Klugman (2001). "Impact of human immunodeficiency virus type 1 infection on the epidemiology and outcome of bacterial meningitis in South African children." Int J Infect Dis **5**(3): 119-125.

Madhi, S. A., B. Schoub, K. Simmank, N. Blackburn and K. P. Klugman (2000). "Increased burden of respiratory viral associated severe lower respiratory tract infections in children infected with human immunodeficiency virus type-1." J Pediatr **137**(1): 78-84.

Marsh, V., D. Kamuya, Y. Rowa, C. Gikonyo and S. Molyneux (2008). "Beginning community engagement at a busy biomedical research programme: experiences from the KEMRI CGMRC-Wellcome Trust Research Programme, Kilifi, Kenya." Soc Sci Med **67**(5): 721-733.

Medley, G. F. and D. J. Nokes (2009). "Epidemiology. Does viral diversity matter?" Science **325**(5938): 274-275.

Meert, K. L., A. P. Sarnaik, M. J. Gelmini and M. W. Lieh-Lai (1994). "Aerosolized ribavirin in mechanically ventilated children with respiratory syncytial virus lower respiratory tract disease: a prospective, double-blind, randomized trial." Crit Care Med **22**(4): 566-572.

Melegaro, A., N. J. Gay and G. F. Medley (2004). "Estimating the transmission parameters of pneumococcal carriage in households." Epidemiol Infect **132**(3): 433-441.

Melero, J., Ed. (2007). Molecular Biology of Human Respiratory Syncytial Virus. Respiratory Syncytial Virus. 2007, Elsevier,.

Mills, J. t., J. E. Van Kirk, P. F. Wright and R. M. Chanock (1971). "Experimental respiratory syncytial virus infection of adults. Possible mechanisms of resistance to infection and illness." J Immunol **107**(1): 123-130.

Milne, G. J., J. K. Kelso, H. A. Kelly, S. T. Huband and J. McVernon (2008). "A small community model for the transmission of infectious diseases: comparison of school closure as an intervention in individual-based models of an influenza pandemic." PLoS One **3**(12): e4005.

Mlinaric-Galinovic, G., T. Chonmaitree, P. A. Cane, C. R. Pringle and P. L. Ogra (1994). "Antigenic diversity of respiratory syncytial virus subgroup B strains circulating during a community outbreak of infection." J Med Virol **42**(4): 380-384.

Mlinaric-Galinovic, G., I. Tabain, T. Kukovec, G. Vojnovic, J. Bozikov, J. Bogovic-Cepin, *et al.* (2012). "Analysis of biennial outbreak pattern of respiratory syncytial virus according to subtype (A and B) in the Zagreb region." Pediatr Int **54**(3): 331-335.

Mlinaric-Galinovic, G., R. C. Welliver, T. Vilibic-Cavlek, S. Ljubin-Sternak, V. Drazenovic, I. Galinovic and V. Tomic (2008). "The biennial cycle of respiratory syncytial virus outbreaks in Croatia." Virol J **5**: 18.

Moisi, J. C., H. Gatakaa, A. M. Noor, T. N. Williams, E. Bauni, B. Tsofa, *et al.* (2010). "Geographic access to care is not a determinant of child mortality in a rural Kenyan setting with high health facility density." BMC Public Health **10**(142): 142.

Moisi, J. C., D. J. Nokes, H. Gatakaa, T. N. Williams, E. Bauni, O. S. Levine and J. A. Scott (2011). "Sensitivity of hospital-based surveillance for severe disease: a geographic information system analysis of access to care in Kilifi district, Kenya." Bull World Health Organ **89**(2): 102-111.

Moler, F. W., C. M. Steinhart, S. E. Ohmit and G. L. Stidham (1996). "Effectiveness of ribavirin in otherwise well infants with respiratory syncytial virus-associated respiratory failure. Pediatric Critical Study Group." J Pediatr **128**(3): 422-428.

Monto, A. S. (2002). "Epidemiology of viral respiratory infections." Am J Med **112 Suppl 6A**: 4S-12S.

Monto, A. S., E. R. Bryan and L. M. Rhodes (1974). "The Tecumseh study of respiratory illness. VII. Further observations on the occurrence of respiratory syncytial virus and Mycoplasma pneumoniae infections." Am J Epidemiol **100**(6): 458-468.

Monto, A. S. and J. J. Cavallaro (1971). "The Tecumseh study of respiratory illness. II. Patterns of occurrence of infection with respiratory pathogens, 1965-1969." Am J Epidemiol **94**(3): 280-289.

Monto, A. S. and J. J. Cavallaro (1972). "The Tecumseh study of respiratory illness. IV. Prevalence of rhinovirus serotypes, 1966-1969." Am J Epidemiol **96**(5): 352-360.

Monto, A. S. and S. K. Lim (1971). "The Tecumseh study of respiratory illness. 3. Incidence and periodicity of respiratory syncytial virus and Mycoplasma pneumoniae infections." Am J Epidemiol **94**(3): 290-301.

Morris, J. A., R. E. Blount, Jr. and R. E. Savage (1956). "Recovery of cytopathogenic agent from chimpanzees with coryza." Proc Soc Exp Biol Med **92**(3): 544-549.

Mossong, J., N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, *et al.* (2008). "Social contacts and mixing patterns relevant to the spread of infectious diseases." PLoS Med **5**(3): e74.

Munoz, F. M., P. A. Piedra and W. P. Glezen (2003). "Safety and immunogenicity of respiratory syncytial virus purified fusion protein-2 vaccine in pregnant women." Vaccine **21**(24): 3465-3467.

Munywoki, P. K., F. Hamid, M. Mutunga, S. Welch, P. Cane and D. J. Nokes (2011). "Improved detection of respiratory viruses in pediatric outpatients with acute respiratory illness by real-time PCR using nasopharyngeal flocked swabs." J Clin Microbiol **49**(9): 3365-3367.

Munywoki, P. K., E. O. Ohuma, M. Ngama, E. Bauni, J. A. Scott and D. J. Nokes (2013). "Severe lower respiratory tract infection in early infancy and pneumonia hospitalizations among children, Kenya." Emerg Infect Dis **19**(2): 223-229.

Murphy, B. R., R. A. Olmsted, P. L. Collins, R. M. Chanock and G. A. Prince (1988). "Passive transfer of respiratory syncytial virus (RSV) antiserum suppresses the immune response to the RSV fusion (F) and large (G) glycoproteins expressed by recombinant vaccinia viruses." J Virol **62**(10): 3907-3910.

Murphy, B. R., A. Sotnikov, P. R. Paradiso, S. W. Hildreth, A. B. Jenson, R. B. Baggs, *et al.* (1989). "Immunization of cotton rats with the fusion (F) and large (G) glycoproteins of respiratory syncytial virus (RSV) protects against RSV challenge without potentiating RSV disease." Vaccine **7**(6): 533-540.

Nair, H., D. J. Nokes, B. D. Gessner, M. Dherani, S. A. Madhi, R. J. Singleton, *et al.* (2010). "Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis." Lancet **375**(9725): 1545-1555.

Ngama, M. J., B. Ouma, M. E. English and D. J. Nokes (2004). "Comparison of three methods of collecting nasal specimens for respiratory virus analysis." East Afr Med J **81**(6): 313-317.

Nokes, D. J. (2007). Respiratory syncytial virus disease burden in the developing world. Perspectives in Medical Virology: Respiratory Syncytial Virus. P. A. Cane. Amsterdam, Elsevier. **14:** 205 -210.

Nokes, D. J., F. Enquselassie, W. Nigatu, A. J. Vyse, B. J. Cohen, D. W. Brown and F. T. Cutts (2001). "Has oral fluid the potential to replace serum for the evaluation of population immunity levels? A study of measles, rubella and hepatitis B in rural Ethiopia." Bull World Health Organ **79**(7): 588-595.

Nokes, D. J., M. Ngama, A. Bett, J. Abwao, P. Munywoki, M. English, *et al.* (2009). "Incidence and severity of respiratory syncytial virus pneumonia in rural Kenyan children identified through hospital surveillance." Clin Infect Dis **49**(9): 1341-1349.

Nokes, D. J., E. A. Okiro, M. Ngama, R. Ochola, L. J. White, P. D. Scott, *et al.* (2008). "Respiratory syncytial virus infection and disease in infants and young children observed from birth in Kilifi District, Kenya." Clin Infect Dis **46**(1): 50-57.

Nokes, D. J., E. A. Okiro, M. Ngama, L. J. White, R. Ochola, P. D. Scott, *et al.* (2004). "Respiratory Syncytial Virus Epidemiology in a Birth Cohort from Kilifi District, Kenya: Infection during the First Year of Life." J Infect Dis **190**(10): 1828-1832.

Nokes, J. D. and P. A. Cane (2008). "New strategies for control of respiratory syncytial virus infection." Curr Opin Infect Dis **21**(6): 639-643.

Ochola, R., C. Sande, G. Fegan, P. D. Scott, G. F. Medley, P. A. Cane and D. J. Nokes (2009). "The level and duration of RSV-specific maternal IgG in infants in Kilifi Kenya." PLoS One **4**(12): e8088.

Ohuma, E. O., E. A. Okiro, R. Ochola, C. J. Sande, P. A. Cane, G. F. Medley, *et al.* (2012). "The natural history of respiratory syncytial virus in a birth cohort: the influence of age and previous infection on reinfection and disease." Am J Epidemiol **176**(9): 794-802.

Okiro, E. A. (2007). Transmission dynamics of respiratory sycyntial virus within the household and in the community. Doctor of Philosphy, University of Warwick.

Okiro, E. A., M. Ngama, A. Bett, P. A. Cane, G. F. Medley and D. James Nokes (2008). "Factors associated with increased risk of progression to respiratory syncytial virus-associated pneumonia in young Kenyan children." Trop Med Int Health **13**(7): 914-926.

Okiro, E. A., C. Sande, M. Mutunga, G. F. Medley, P. A. Cane and D. J. Nokes (2008). "Identifying infections with respiratory syncytial virus by using specific immunoglobulin G (IgG) and IgA enzyme-linked immunosorbent assays with oral-fluid samples." J Clin Microbiol **46**(5): 1659-1662.

Okiro, E. A., L. J. White, M. Ngama, P. A. Cane, G. F. Medley and D. J. Nokes (2010). "Duration of shedding of respiratory syncytial virus in a community study of Kenyan children." BMC Infect Dis **10**(15): 15.

Olsen, S. J., S. Thamthitiwat, S. Chantra, M. Chittaganpitch, A. M. Fry, J. M. Simmerman, *et al.* (2010). "Incidence of respiratory pathogens in persons hospitalized with pneumonia in two provinces in Thailand." Epidemiol Infect **138**(12): 1811-1822.

Paradiso, P. R., S. W. Hildreth, D. A. Hogerman, D. J. Speelman, E. B. Lewin, J. Oren and D. H. Smith (1994). "Safety and immunogenicity of a subunit respiratory syncytial virus vaccine in children 24 to 48 months old." Pediatr Infect Dis J **13**(9): 792-798.

Phillips, P. A., D. Lehmann, V. Spooner, J. Barker, S. Tulloch, M. Sungu, *et al.* (1990). "Viruses associated with acute lower respiratory tract infections in children from the eastern highlands of Papua New Guinea (1983-1985)." Southeast Asian J Trop Med Public Health **21**(3): 373-382.

Piedra, P. A., S. G. Cron, A. Jewell, N. Hamblett, R. McBride, M. A. Palacio, *et al.* (2003). "Immunogenicity of a new purified fusion protein vaccine to respiratory syncytial virus: a multi-center trial in children with cystic fibrosis." Vaccine **21**(19-20): 2448-2460.

Piedra, P. A., A. M. Jewell, S. G. Cron, R. L. Atmar and W. P. Glezen (2003). "Correlates of immunity to respiratory syncytial virus (RSV) associated-hospitalization: establishment of minimum protective threshold levels of serum neutralizing antibodies." Vaccine **21**(24): 3479-3482.

Potter, G. E., M. S. Handcock, I. M. Longini, Jr. and M. E. Halloran (2011). "Estimating within-Household Contact Networks from Egocentric Data." Ann Appl Stat **5**(3): 1816-1838.

Poulsen, A., L. G. Stensballe, J. Nielsen, C. S. Benn, A. Balde, A. Roth, *et al.* (2006). "Long-term consequences of respiratory syncytial virus acute lower respiratory tract infection in early childhood in Guinea-bissau." Pediatr Infect Dis J **25**(11): 1025-1031.

Power, U. F., T. N. Nguyen, E. Rietveld, R. L. de Swart, J. Groen, A. D. Osterhaus, *et al.* (2001). "Safety and immunogenicity of a novel recombinant subunit respiratory syncytial virus vaccine (BBG2Na) in healthy young adults." J Infect Dis **184**(11): 1456-1460.

Power, U. F., H. Plotnicky, A. Blaecke and T. N. Nguyen (2003). "The immunogenicity, protective efficacy and safety of BBG2Na, a subunit respiratory syncytial virus (RSV) vaccine candidate, against RSV-B." Vaccine **22**(2): 168-176.

R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Read, J. M., K. T. Eames and W. J. Edmunds (2008). "Dynamic social networks and the implications for the spread of infectious disease." J R Soc Interface **5**(26): 1001-1007.

Reese, P. E. and N. J. Marchette (1991). "Respiratory syncytial virus infection and prevalence of subgroups A and B in Hawaii." J Clin Microbiol **29**(11): 2614-2615.

Robertson, S. E., A. Roca, P. Alonso, E. A. Simoes, C. B. Kartasasmita, D. O. Olaleye, *et al.* (2004). "Respiratory syncytial virus infection: denominator-based studies in Indonesia, Mozambique, Nigeria and South Africa." Bull World Health Organ **82**(12): 914-922.

Robinson, J. L., B. E. Lee, S. Kothapalli, W. R. Craig and J. D. Fox (2008). "Use of throat swab or saliva specimens for detection of respiratory viruses in children." Clin Infect Dis **46**(7): e61-64.

Roca, A., M. P. Loscertales, L. Quinto, P. Perez-Brena, N. Vaz, P. L. Alonso and J. C. Saiz (2001). "Genetic variability among group A and B respiratory syncytial viruses in Mozambique: identification of a new cluster of group B isolates." J Gen Virol **82**(Pt 1): 103-111.

Rodriguez, W. J. and R. H. Parrott (1987). "Ribavirin aerosol treatment of serious respiratory syncytial virus infection in infants." Infect Dis Clin North Am **1**(2): 425-439.

Rudan, I., E. Theodoratou, L. Zgaga, H. Nair, K. Y. Chan, M. Tomlinson, *et al.* (2012). "Setting priorities for development of emerging interventions against childhood pneumonia, meningitis and influenza." J Glob Health **2**(1): 10304.

Sakthivel, S. K., B. Whitaker, X. Lu, D. B. Oliveira, L. J. Stockman, S. Kamili, *et al.* (2012). "Comparison of fast-track diagnostics respiratory pathogens multiplex real-time RT-PCR assay with in-house singleplex assays for comprehensive detection of human respiratory viruses." J Virol Methods **185**(2): 259-266.

Sande, C. J., M. N. Mutunga, G. F. Medley, P. A. Cane and D. J. Nokes (2013). "Group- and genotype-specific neutralizing antibody responses against respiratory syncytial virus in infants and young children with severe pneumonia." J Infect Dis **207**(3): 489-492.

Schenzle, D. (1984). "An age-structured model of pre- and post-vaccination measles transmission." IMA Journal of Mathematics Applied in Medicine & Biology **1**: 169-191.

Schmidt, A. C., J. M. McAuliffe, B. R. Murphy and P. L. Collins (2001). "Recombinant bovine/human parainfluenza virus type 3 (B/HPIV3) expressing the respiratory

syncytial virus (RSV) G and F proteins can be used to achieve simultaneous mucosal immunization against RSV and HPIV3." J Virol **75**(10): 4594-4603.

Schwarze, J., D. R. O'Donnell, A. Rohwedder and P. J. Openshaw (2004). "Latency and persistence of respiratory syncytial virus despite T cell immunity." Am J Respir Crit Care Med **169**(7): 801-805.

Scott, J. A., E. Bauni, J. C. Moisi, J. Ojal, H. Gatakaa, C. Nyundo, *et al.* (2012). "Profile: The Kilifi Health and Demographic Surveillance System (KHDSS)." Int J Epidemiol **41**(3): 650-657.

Scott, P. D., R. Ochola, M. Ngama, E. A. Okiro, D. James Nokes, G. F. Medley and P. A. Cane (2006). "Molecular analysis of respiratory syncytial virus reinfections in infants from coastal Kenya." J Infect Dis **193**(1): 59-67.

Scott, P. D., R. Ochola, M. Ngama, E. A. Okiro, D. J. Nokes, G. F. Medley and P. A. Cane (2004). "Molecular epidemiology of respiratory syncytial virus in Kilifi district, Kenya." J Med Virol **74**(2): 344-354.

Scott, P. D., R. Ochola, C. Sande, M. Ngama, E. A. Okiro, G. F. Medley, *et al.* (2007). "Comparison of strain-specific antibody responses during primary and secondary infections with respiratory syncytial virus." J Med Virol **79**(12): 1943-1950.

Seki, K., H. Tsutsumi, M. Ohsaki, H. Kamasaki and S. Chiba (2001). "Genetic variability of respiratory syncytial virus subgroup a strain in 15 successive epidemics in one city." J Med Virol **64**(3): 374-380.

Semple, M. G., A. Cowell, W. Dove, J. Greensill, P. S. McNamara, C. Halfhide, *et al.* (2005). "Dual infection of infants by human metapneumovirus and human respiratory syncytial virus is strongly associated with severe bronchiolitis." J Infect Dis **191**(3): 382-386.

Shanno, D. F. (1970). "Conditioning of quasi-Newton methods for function minimization." Mathematics of computation **24**(111): 647-656.

Simoes, E. A., H. M. Sondheimer, F. H. Top, Jr., H. C. Meissner, R. C. Welliver, A. A. Kramer and J. R. Groothuis (1998). "Respiratory syncytial virus immune globulin for prophylaxis against respiratory syncytial virus disease in infants and children with congenital heart disease. The Cardiac Study Group." J Pediatr **133**(4): 492-499.

Simoes, E. A. F. (1999). "Respiratory syncytial virus infection." Lancet **354**(9181): 847-852.

Sims, D. G., M. A. Downham, J. McQuillin and P. S. Gardner (1976). "Respiratory syncytial virus infection in north-east England." Br Med J **2**(6044): 1095-1098.

Singleton, R. J., D. Bruden, L. R. Bulkow, G. Varney and J. C. Butler (2006). "Decline in respiratory syncytial virus hospitalizations in a region with high hospitalization rates and prolonged season." Pediatr Infect Dis J **25**(12): 1116-1122.

Smith, D. W., L. R. Frankel, L. H. Mathers, A. T. Tang, R. L. Ariagno and C. G. Prober (1991). "A controlled trial of aerosolized ribavirin in infants receiving mechanical ventilation for severe respiratory syncytial virus infection." N Engl J Med **325**(1): 24-29.

Smith, P. G. and R. H. Morrow (1992). Methods for Field Trials of Interventions Against Tropical Diseases: a 'Toolbox', Oxford University Press.

Souza, L. S., E. A. Ramos, F. M. Carvalho, V. M. Guedes, C. M. Rocha, A. B. Soares, *et al.* (2003). "Viral respiratory infections in young children attending day care in urban Northeast Brazil." Pediatr Pulmonol **35**(3): 184-191.

Stensballe, L., A. Poulsen, E. Nante, I. P. Jensen, P. E. Kofoed, H. Jensen and P. Aaby (2004). "Mothers may transmit RSV infection more easily or severely to sons than daughters: community study from Guinea-Bissau." Scand J Infect Dis **36**(4): 291-295.

Stensballe, L. G., J. K. Devasundaram and E. A. Simoes (2003). "Respiratory syncytial virus epidemics: the ups and downs of a seasonal virus." Pediatr Infect Dis J **22**(2 Suppl): S21-32.

Stockton, J., J. S. Ellis, M. Saville, J. P. Clewley and M. C. Zambon (1998). "Multiplex PCR for typing and subtyping influenza and respiratory syncytial viruses." J Clin Microbiol **36**(10): 2990-2995.

Sutmoller, F., Z. P. Ferro, M. D. Asensi, V. Ferreira, I. S. Mazzei and B. L. Cunha (1995). "Etiology of acute respiratory tract infections among children in a combined community and hospital study in Rio de Janeiro." Clin Infect Dis **20**(4): 854-860.

Suwanjutha, S., P. Sunakorn, T. Chantarojanasiri, S. Siritantikorn, S. Nawanoparatkul, T. Rattanadilok Na Bhuket, *et al.* (2002). "Respiratory syncytial virus-associated lower respiratory tract infection in under-5-year-old children in a rural community of central Thailand, a population-based study." J Med Assoc Thai **85 Suppl 4**: S1111-1119.

Taber, L. H., V. Knight, B. E. Gilbert, H. W. McClung, S. Z. Wilson, H. J. Norton, *et al.* (1983). "Ribavirin aerosol treatment of bronchiolitis associated with respiratory syncytial virus infection in infants." Pediatrics **72**(5): 613-618.

Taylor, G., Ed. (2007). Immunology of RSV. Respiratory Syncytial Virus. 2007, Elsevier,.

Taylor, G., E. J. Stott, M. Bew, B. F. Fernie, P. J. Cote, A. P. Collins, *et al.* (1984). "Monoclonal antibodies protect against respiratory syncytial virus infection in mice." Immunology **52**(1): 137-142.

Taylor, G., E. J. Stott, J. Furze, J. Ford and P. Sopp (1992). "Protective epitopes on the fusion protein of respiratory syncytial virus recognized by murine and bovine monoclonal antibodies." J Gen Virol **73 (Pt 9)**: 2217-2223.

Teng, M. N., S. S. Whitehead, A. Bermingham, M. St Claire, W. R. Elkins, B. R. Murphy and P. L. Collins (2000). "Recombinant respiratory syncytial virus that does not

express the NS1 or M2-2 protein is highly attenuated and immunogenic in chimpanzees." J Virol **74**(19): 9317-9321.

Thompson, W. W., D. K. Shay, E. Weintraub, L. Brammer, N. Cox, L. J. Anderson and K. Fukuda (2003). "Mortality associated with influenza and respiratory syncytial virus in the United States." JAMA **289**(2): 179-186.

Trento, A., I. Casas, A. Calderon, M. L. Garcia-Garcia, C. Calvo, P. Perez-Brena and J. A. Melero (2010). "Ten years of global evolution of the human respiratory syncytial virus BA genotype with a 60-nucleotide duplication in the G protein gene." J Virol **84**(15): 7500-7512.

Trento, A., M. Galiano, C. Videla, G. Carballal, B. Garcia-Barreno, J. A. Melero and C. Palomo (2003). "Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides." J Gen Virol **84**(Pt 11): 3115-3120.

Trento, A., M. Viegas, M. Galiano, C. Videla, G. Carballal, A. S. Mistchenko and J. A. Melero (2006). "Natural history of human respiratory syncytial virus inferred from phylogenetic analysis of the attachment (G) glycoprotein with a 60-nucleotide duplication." J Virol **80**(2): 975-984.

Tupasi, T. E., L. E. de Leon, S. Lupisan, C. U. Torres, Z. A. Leonor, E. S. Sunico, *et al.* (1990). "Patterns of acute respiratory tract infection in children: a longitudinal study in a depressed community in Metro Manila." Rev Infect Dis **12 Suppl 8**: S940-949.

UNICEF, WHO, World Bank and UN (2012). Levels and trends in child mortality: report 2012. New York.

van den Hoogen, B. G., G. J. van Doornum, J. C. Fockens, J. J. Cornelissen, W. E. Beyer, R. de Groot, *et al.* (2003). "Prevalence and clinical symptoms of human metapneumovirus infection in hospitalized patients." J Infect Dis **188**(10): 1571-1577.

van Elden, L. J., A. M. van Loon, A. van der Beek, K. A. Hendriksen, A. I. Hoepelman, M. G. van Kraaij, *et al.* (2003). "Applicability of a real-time quantitative PCR assay for diagnosis of respiratory syncytial virus infection in immunocompromised adults." J Clin Microbiol **41**(9): 4378-4381.

Vardas, E., D. Blaauw and J. McAnerney (1999). "The epidemiology of respiratory syncytial virus (RSV) infections in South African children." S Afr Med J **89**(10): 1079-1084.

Venter, M., M. Collinson and B. D. Schoub (2002). "Molecular epidemiological analysis of community circulating respiratory syncytial virus in rural South Africa: Comparison of viruses and genotypes responsible for different disease manifestations." J Med Virol **68**(3): 452-461.

Ventre, K. and A. G. Randolph (2007). "Ribavirin for respiratory syncytial virus infection of the lower respiratory tract in infants and young children." Cochrane Database Syst Rev(1): CD000181.

Viboud, C., P. Y. Boelle, S. Cauchemez, A. Lavenu, A. J. Valleron, A. Flahault and F. Carrat (2004). "Risk factors of influenza transmission in households." Br J Gen Pract **54**(506): 684-689.

von Linstow, M. L., J. Eugen-Olsen, A. Koch, T. N. Winther, H. Westh and B. Hogh (2006). "Excretion patterns of human metapneumovirus and respiratory syncytial virus among young children." Eur J Med Res **11**(8): 329-335.

Vynnycky, E., R. Pitman, R. Siddiqui, N. Gay and W. J. Edmunds (2008). "Estimating the impact of childhood influenza vaccination programmes in England and Wales." Vaccine **26**(41): 5321-5330.

Walsh, E. E., J. J. Schlesinger and M. W. Brandriss (1984). "Protection from respiratory syncytial virus infection in cotton rats by passive transfer of monoclonal antibodies." Infect Immun **43**(2): 756-758.

Wang, E. E., B. J. Law and D. Stephens (1995). "Pediatric Investigators Collaborative Network on Infections in Canada (PICNIC) prospective study of risk factors and outcomes in patients hospitalized with respiratory syncytial viral lower respiratory tract infection." J Pediatr **126**(2): 212-219.

Waris, M. (1991). "Pattern of respiratory syncytial virus epidemics in Finland: two-year cycles with alternating prevalence of groups A and B." J Infect Dis **163**(3): 464-469.

Waris, M., O. Meurman, M. A. Mufson, O. Ruuskanen and P. Halonen (1992). "Shedding of infectious virus and virus antigen during acute infection with respiratory syncytial virus." J Med Virol **38**(2): 111-116.

Weber, M. W., R. Dackour, S. Usen, G. Schneider, R. A. Adegbola, P. Cane, *et al.* (1998). "The clinical spectrum of respiratory syncytial virus disease in The Gambia." Pediatric infectious disease journal **17**(3): 224-230.

Weber, M. W., P. Milligan, B. Giadom, M. A. Pate, A. Kwara, A. D. Sadiq, *et al.* (1999). "Respiratory illness after severe respiratory syncytial virus disease in infancy in The Gambia." J Pediatr **135**(6): 683-688.

Weber, M. W., P. Milligan, M. Sanneh, A. Awemoyi, R. Dakour, G. Schneider, *et al.* (2002). "An epidemiological study of RSV infection in the Gambia." Bull World Health Organ **80**(7): 562-568.

Weber, M. W., E. K. Mulholland and B. M. Greenwood (1998). "Respiratory syncytial virus infection in tropical and developing countries." Tropical Medicine and Internationl Health **3**(4): 268-280.

West, K., J. Bogdan, A. Hamel, G. Nayar, P. S. Morley, D. M. Haines and J. A. Ellis (1998). "A comparison of diagnostic methods for the detection of bovine respiratory syncytial virus in experimental clinical specimens." Can J Vet Res **62**(4): 245-250.

White, L. J., J. N. Mandl, M. G. Gomes, A. T. Bodley-Tickell, P. A. Cane, P. Perez-Brena, *et al.* (2007). "Understanding the transmission dynamics of respiratory syncytial virus using multiple time series and nested models." Math Biosci **209**(1): 222-239.

White, L. J., M. Waris, P. A. Cane, D. J. Nokes and G. F. Medley (2005). "The transmission dynamics of groups A and B human respiratory syncytial virus (hRSV) in England & Wales and Finland: seasonality and cross-protection." Epidemiol Infect **133**(2): 279-289.

WHO (2006) "Collecting, preserving and shipping specimens for the diagnosis of avian influenza A(H5N1) virus infection: Guide for field operations." Viral transport media (VTM), 42 DOI: WHO/CDS/EPR/ARO/2006.1.

Widjojoatmodjo, M. N., J. Boes, M. van Bers, Y. van Remmerden, P. J. Roholl and W. Luytjes (2010). "A highly attenuated recombinant human respiratory syncytial virus lacking the G protein induces long-lasting protection in cotton rats." Virol J **7**: 114.

Wilson, J. N., D. J. Nokes, G. F. Medley and D. Shouval (2007). "Mathematical model of the antibody response to hepatitis B vaccines: implications for reduced schedules." Vaccine **25**(18): 3705-3712.

Wright, P. F., R. A. Karron, R. B. Belshe, J. R. Shi, V. B. Randolph, P. L. Collins, *et al.* (2007). "The absence of enhanced disease with wild type respiratory syncytial virus infection occurring after receipt of live, attenuated, respiratory syncytial virus vaccines." Vaccine **25**(42): 7372-7378.

Wright, P. F., R. A. Karron, R. B. Belshe, J. Thompson, J. E. Crowe, Jr., T. G. Boyce, *et al.* (2000). "Evaluation of a live, cold-passaged, temperature-sensitive, respiratory syncytial virus vaccine candidate in infancy." J Infect Dis **182**(5): 1331-1342.

Zagheni, E., F. C. Billari, P. Manfredi, A. Melegaro, J. Mossong and W. J. Edmunds (2008). "Using time-use data to parameterize models for the spread of close-contact infectious diseases." Am J Epidemiol **168**(9): 1082-1090.

Zambon, M. C., J. D. Stockton, J. P. Clewley and D. M. Fleming (2001). "Contribution of influenza and respiratory syncytial virus to community cases of influenza-like illness: an observational study." Lancet **358**(9291): 1410-1416.

Zlateva, K. T., P. Lemey, E. Moes, A. M. Vandamme and M. Van Ranst (2005). "Genetic variability and molecular evolution of the human respiratory syncytial virus subgroup B attachment G protein." J Virol **79**(14): 9157-9167.

Zlateva, K. T., L. Vijgen, N. Dekeersmaeker, C. Naranjo and M. Van Ranst (2007). "Subgroup prevalence and genotype circulation patterns of human respiratory syncytial virus in Belgium during ten successive epidemic seasons." J Clin Microbiol **45**(9): 3022-3030.

10 APPENDICES

A number of appendices are attached. These attachments provide further details on study starting with the study proposal, scientific and ethical approval letters, field worker training, consent documents, data collection tools, and standard operating procedures on sample collection and testing. Additional results are also included in the appendix such as on RSV infection patterns within the 47 households, estimates of duration of RSV shedding, recovery rates and unadjusted hazard ratios for factors likely to influence cessation of RSV shedding. Finally, the R code used in estimation of susceptibility and transmission parameters is also attached.

Appendix A. The study Protocol

a) Title of the Project

Household transmission of respiratory syncytial virus (RSV): Who acquires infection from whom

b) Investigators and Institutional Affiliations

Name	Institution
Principal Investigator	
Patrick Munywoki	CGMRC
Co-Investigators	
Dr. James Nokes	CGMR-C, University of Warwick, UK
Mwanajuma Ngama	CGMR-C
Charles Nyaigoti	CGMR-C
Ann Bett	CGMR-C
Dorothy Koech	CGMR-C
Dr Laura Hammit	CGMR-C
Dr Patricia Cane	Health Protection Agency, UK
Prof. Graham Medley	University of Warwick, UK

c) Abstract

Respiratory syncytial virus (RSV) is a major cause of childhood acute respiratory infection worldwide. Studies from Kilifi District show that around 1 in every 100 infants are admitted to hospital with RSV associated severe pneumonia each year. A vaccine is not yet available to prevent RSV infection but is under development. The potential usefulness of such a vaccine and information on how best it can be implemented requires a better understanding of the spread of the virus within the community. The family or household unit is an important element in the spread of infections transmitted by close contact in the community and better understanding of household transmission may inform on the merits of targeted control strategies. Data is scarce on RSV transmission within families, particularly on who acquires infection from whom (WAIFW). Mathematical modelling of infection transmission has been previously used to evaluate the impact of different population-based vaccination strategies. However, for such models to be realistic requires well-determined parameters on WAIFW and especially who is infecting the vulnerable infants.

With an aim of elucidating who infects the infant within the household, we intend to recruit 50 households prior to, and undertake intensive follow-up during a RSV epidemic. Households will be eligible if they have a child born after the previous RSV epidemic and with at least one elder sibling to the infant. Nasal samples will be taken every 3-4 days from all household members, irrespective of symptoms, for 4-5 months (while the RSV epidemic persists). Oral fluid will be collected once a week. Samples will be screened using PCR methods for a range of respiratory viruses including RSV and other co-circulating viruses with similar transmission routes such as rhinovirus, human metapneumovirus (hMPV) and

parainfluenza virus 3 (PIV 3). Saliva samples will be screened using the antibody capture ELISA methods and directly for viral antigens by molecular methods. The detailed multiple infection data, arising from the study, will be used to define and quantify who infects whom within the household, and especially who is the source of the virus that infects the infant. Understanding of transmission dynamics of RSV will help elucidate the potential of different vaccination strategies and the impact at the population level of such approaches on RSV infection especially on the vulnerable infants.

d) Introduction/Background

RSV disease burden and control

RSV is a major cause of childhood acute respiratory infection (ARI) worldwide. It is estimated that about 60% of all children are infected during their first year of life and almost all by their third year (Glezen et al. 1986). Our surveillance of paediatric admissions in rural coastal Kenya, (Kilifi district), over the period 2002-07 has defined incidence estimates (per 100,000 per year) of RSV-associated severe or very severe pneumonia admissions of 1107 (95% CI, 1012-1211) in infants, and 293 (271-317) in the under 5 year olds, contributing some 15% and 12%, by age group, respectively, of all admissions for these conditions (Nokes DJ et al 2009). This is comparable with the proportion of clinical severe pneumonia prevented by 9-valent conjugate pneumococcal vaccine in The Gambia (12%)(Cutts et al. 2005). Spatial analysis suggests that the true incidence of RSV pneumonias is well in excess of these estimates (due to distance effects on health service access).

RSV repeatedly reinfects throughout life (Hall et al. 1976; Henderson et al. 1979). Undoubtedly this translates to a large pool of infectious individuals likely to be fundamental to endemic maintenance of infection and transmission to young children. The contribution of re-infections to the spread of RSV within the community is not well elucidated, but will be dependent upon the prevalence of individuals with re-infections, and their infectivity and patterns of contact with others in the population. These factors combine to define 'who acquires infection from whom' (WAIFW), which is central to understanding the transmission dynamics, and predicting vaccine impact (Anderson and May 1991).

There is a body of evidence which demonstrates that viral infection of respiratory epithelial cells promotes bacterial colonization of the nasopharynx. RSV infection is associated with secondary bacterial infections caused by Streptococcus pneumonia and Haemophilus influenza (Avadhanula et al. 2006; Pettigrew et al. 2008). As such, the burden of RSV-related disease may be much higher than is estimated by the number of RSV infections alone.

Currently there is no licensed RSV vaccine but promising live-attenuated vaccines are under development (Karron et al. 2005). With most severe RSV disease occurring in the first year of life, and especially in the first 6 months, it is a priority that a vaccine protects this early age group. A highly attenuated recombinant vaccine variant for delivery to infants under 3 months of age, was found to be well tolerated, but was insufficiently immunogenic, although in children 6 months of age and older (naïve and seropositive) it was well tolerated, and immunogenic (Karron et al. 2005; Wright et al. 2007). The potential of such vaccines to protect the most vulnerable children depends not only on the direct protection to the

vaccinated (≥3 month old) individuals but also on resultant reduced spread of infection from those vaccinated to those too young to receive vaccine (<3 months). This indirect protection may be in the form of a reduced risk of transmission between presumed key transmission contacts (for example, mother to infant), or more generally due to reduced circulation of the infectious agent in the population, the so called herd immunity effect (Anderson and May 1991). These indirect protective or herd immunity effects are largely unquantified. Mathematical modelling of infection transmission has been previously used to evaluate the impact of different vaccination strategies (Anderson and Grenfell 1986; Babad et al. 1995; Vynnycky et al. 2008). However, for such models to be realistic requires well-determined parameters for WAIFW and, in the case of RSV, who is infecting vulnerable infants.

Other respiratory viruses

The advent of molecular diagnostics for a broad range of respiratory viruses has enhanced the study of virus epidemiology, due to increased sensitivity and range of pathogens detectable over traditional methods. Throughout 2007 we used molecular methods to investigate the occurrence of 13 respiratory viruses in 760 severe and very severe pneumonia paediatric admissions to Kilifi District Hospital (Dr. Jay Berkley, unpublished data, SSC no. 815). RSV was detected in 34% of cases (compared to only 5% of 56 well controls) using molecular methods, compared to 21% of cases by immunofluorescence. At least one virus was present in 56% of cases, with peak occurrence in the months of November, December and January to March, coincident with peak pneumonia admissions. Other respiratory viruses have the same mode of transmission as RSV via large droplets and fomites. In the study described above we found co-occurrence temporally of hMPV and PIV3 in particular. Furthermore, rhinovirus is elsewhere reported to be prevalent throughout much of the year and not uncommonly co-circulating with RSV (Phillips et al. 1990; Monto 2002; Souza et al. 2003; Kusel et al. 2006). Screening for viruses with similar epidemiology (such as rhinovirus, hMPV and PIV3) to RSV would increase the capacity of the present study to identify infection spread in the household and hence improve the definition and quantification of WAIFW for a given sample size.

Who contacts whom and who infects whom

RSV is transmitted by large nasal droplets and fomites, with short survival time in the environment. This suggests that close contact is important for effective transmission. Social interaction patterns (contacts and mixing patterns among individuals) have thus a direct impact on transmission dynamics of respiratory infections. This underlies the growing interest in describing and quantifying contacts that can lead to infection spread (Edmunds et al. 1997; Edmunds et al. 2006; Mossong et al. 2008; Read et al. 2008; Vynnycky et al. 2008; Zagheni et al. 2008). The use of contact data to define infectious contact rates in transmission dynamics modelling requires inference of the transmission probability following contact – which is a serious obstacle due to its considerable uncertainty. Longitudinal studies of infection and contact can estimate this transmission probability more directly (Melegaro et al. 2004).

It is plausible that social contexts with high contact rates such as households and schools would provide very favourable environments for infection transmission. Family studies indicate the importance of household size and number of school-age siblings as risk factors of RSV infection in infants, and partly explain re-infection rates (Hall et al. 1976; Okiro 2007). Various studies indicate RSV infection is frequently introduced to the home by siblings, resulting in high secondary attack rates within households (Berglund 1967; Hall et al. 1976; Okiro 2007). However a recent study has had contrasting findings, reporting little evidence for infection from sibling in families (Crowcroft et al. 2008). Direct evidence of mother-to-child infection has been reported in one study from Guinea Bissau (Stensballe et al. 2004). It remains, however, that patterns of transmission, so important to determine the influence of infection (or control) in one age group on other groups, are poorly defined.

Household studies

The family or household unit is an important element in the spread of infectious disease in communities, especially where transmission is effected by close contact, and has provided the focus of observational studies in the past. Better understanding of household transmission may inform on the merits of targeted control strategies. For example, investigation of who introduces infection into the household, and who infects the infant, could identify key groups for vaccination such as elder and school age siblings. Recent data is scarce on intra-family spread of RSV and other viral respiratory infections especially in the developing countries. Of the few studies undertaken, most were conducted in 1960s - 70s in developed countries (Monto et al. 1971; Fox and Hall 1972; Hall et al. 1976), with only a single exception in a developing country (Okiro 2007). However the Okiro study had limitations as explained later. The classic family study on RSV spread recruited 36 US families for intensive surveillance over two months in one RSV season, with repeated nasal sampling every 3-4 days regardless of illness (Hall et al. 1976). The authors reported an appreciable secondary attack rate within the family of 27%, rising to 45% in infants.

The study conducted by Okiro et al in our settings aimed at investigating the risk factors of RSV spread and severity and thus its methodology had limitations for the purpose of evaluating WAIFW in the household (Okiro 2007; Okiro et al. 2008; Okiro et al. 2008). Home visits were conducted once-a-week with nasal washings (NW) collected only when symptoms of ARI were reported. Acceptance of the NW method in older children and especially in adults was poor. Screening was by immunofluorescence antigen test (IFAT), which is less sensitive than molecular methods (Casiano-Colon et al. 2003).

We plan to adopt an alternative nasal specimen collection method, the nasopharyngeal flocked swab, (Chan et al. 2008)) which our studies (SSC No. 1527) have shown to be more acceptable without loss of sensitivity compared to nasal washing (Patrick Munywoki et al, unpublished data). Reliance on the occurrence of symptoms to identify infected persons will almost certainly result in failure to recognize links in transmission events. Effort is needed to detect all RSV infection which may frequently be mild, of short duration or sub-clinical in older children and adults (Henderson et al. 1979). Given a mean duration of RSV shedding of between 3.5 and 9 days (dependent upon age, severity, mode of collection and method of antigen detection)(Hall et al. 1976; von Linstow et al. 2006; Okiro 2007), sampling intervals

less frequent than 2 times per week will result in a significant loss in cases detectable (see protocol methods). Recent studies in Kilifi use intensive nasopharyngeal swabbing, including newborns, more than once a week, similar to that proposed in the present study (SSC nos. 782, 1048 and 1071).

The study of Okiro et al collected oral fluid samples to detect specific anti-RSV antibodies to supplement infection data. However, we now know the sampling interval was too infrequent (once every 3 months) to be able to identify infections at the required temporal resolution based on antibody profiles (Okiro et al. 2008). Our recent work suggests that levels of RSV specific IgG (and to a lesser extent IgA) in OF track those of serum, but with very rapid post-infection antibody dynamics (Okiro et al. 2008), suggesting that weekly sampling could help in identifying of infections. Emerging evidence indicates OF samples can be used to detect viral infections using PCR methods but with relatively low sensitivity, and this will specifically be explored (von Linstow et al. 2006).

With the chief aim of defining "Who Acquires Infection From Whom", we plan to mitigate the above limitations by using the most favourable infection detection methods. This will include frequent nasal (twice-a-week) and OF (once-a-week) sampling irrespective of symptoms, and use of highly sensitive assays for identification of viruses. Antibody profiling of OF will provide additional data on infection status. No previous study has combined all these characteristics. Furthermore, the use of PCR based methods to genotype the infecting strain will provide an added benefit in offering a degree of finger printing of clusters of infection.

e) Justification for the Study

Infants are highly vulnerable to severe RSV pneumonia and/or bronchiolitis, and are of crucial importance to control efforts, but no detailed data exists on who infects the infant. Such data would be useful in considering targeted vaccination options. More generally, the role of the household in maintaining transmission of this important virus in the community is not known. Only detailed studies of the intensity proposed will inform our understanding of who is infecting whom and the population-level impact of various vaccination strategies. Twice-a-week nasopharyngeal sampling and once-a-week oral fluid collection, irrespective of symptoms, is critical to ensuring all RSV (and other respiratory viruses) infections occurring in household are identified. Frequent nasal swabbing has been successfully used in Kilifi for pneumococcal carriage studies (SSC nos. 1071, 1048 and 782) and elsewhere (Hall et al. 1976). Use of more sensitive viral screening methods (PCR) is expected to significantly improve detection of RSV and other respiratory pathogens.

f) State the Null Hypothesis

The study is descriptive in nature and has no primary hypothesis.

g) Objectives

General objective

To develop understanding of who acquires infection from whom in the context of introduction and spread of RSV within the household in a developing country setting

Primary objectives

i) To determine the proportion of primary RSV infections in infants that arises from infection from elder siblings or from mothers.
ii) To quantify the rate of infection and recovery of RSV in the family, stratified by age group.
iii) To estimate the household reproductive potential of RSV and risk of transmission per contact.

Secondary objectives

i) To evaluate the use of oral fluids in detecting viral infections using PCR methods.
ii) To estimate the rate of infection and recovery of a range of respiratory viruses for various age groups in the family
iii) To evaluate the effect of viral infection on bacterial colonization of the nasopharynx.

h) Design and Methodology

Study site (geographical)

This study will be conducted within Kilifi District on the Kenyan coast. A suitable location will be selected through an ongoing consultative engagement between the Ministry of health and KEMRI/Wellcome Trust Research Programme.

Study population

The basic study unit will be households defined as people eating from the same kitchen regardless of their relationships.

Criteria for inclusion of households

Households:

i) Within a pre-defined location of Kilifi District
ii) With a child born after the previous RSV epidemic and with at least one elder sibling to the infant (minimum number per household 3 i.e. infant, sibling and parent).
iii) Willingness (including consent) of household to participate in full knowledge of sampling regime and method (demonstrated during the early phase of the study).

Criteria for exclusion of households

i) Failure to get individual consent from all the family members
ii) Household planning to move out of the study area in the next 4 months.

Sampling

Sample size determination

The sample size was set at 50 households on the basis of past evidence on household infection rates and of practical considerations of workload.

Based on previous studies we expect 60% or over of families to be infected during the epidemic. A sample of 50 will give adequate precision in estimates of the prevalence of key outcomes: (i) households with one occupant infected with RSV at the end of the study, (ii) infants experiencing a primary RSV infection and (iii) infants being infected from within (elder siblings or parents) and from outside the immediate household. For the sample size of 50, a proportion of households having an RSV infection of 60% will have 95% confidence limits of 45% and 74% based on binomial exact methods.

Sampling frequency

The following rationale was used to determine the twice-weekly (every 3.5 days) nasal sampling frequency. Assuming individuals shed virus with mean duration of between 3.5 and 9 days, with a constant rate of recovery from shedding, and an onset on average half way between any sampling interval, then the proportion of individuals predicted to remain shedding, and thus detectable, will range from 61%-82% (for 3.5-9 days duration) for a 3.5 day sampling interval (Figure in Appendix B). The comparable range is 37%-68% for a 7-day interval. Given the need to detect infection in mild cases and in older children and adults with likely lower range of shedding duration, sampling twice weekly is indicated.

Sampling procedures

A quick mapping exercise will be carried out to identify eligible households within the selected study location before the start of forthcoming RSV epidemic (in October 2009). Specimen collection will be initiated in November 2009. The first 4 weeks will allow time for the study team to develop a logistical framework/system synchronised with the community activities and culture, and replace non-compliant households, before the epidemic starts. We will use the KDH inpatient RSV surveillance (SSC No. 1055) data to determine the beginning of the RSV season, defined as two successive weeks with > 2 RSV infections (Nokes et al. 2004). Nasopharyngeal flocked swabs (NPS) will be used to collect specimens every 3-4 days from all household members, irrespective of symptoms, throughout the RSV epidemic. During the twice-a-week home visits data on history of respiratory illness will be taken and oral fluid sample collected on one of the visits. Data on a range of household characteristics (e.g. income, education level etc) will also be collected after one month of surveillance, after a good relationship with the family has been established. Follow-up will be terminated for an individual household when its members repeatedly fail to comply with sampling. The study will be ended for all households when the RSV epidemic ends (≤2 RSV infections detected from the KDH inpatient admission samples on two successive weeks).

Procedures

Clinical

Preliminary data from a study evaluating the diagnostic performance of the nasal flocked swab as a specimen option for RSV screening in the community (SSC No. 1527) indicate it is has similar diagnostic performance and more acceptable than our routine method of nasal washing. A brief description of the two specimen collection methods follows.

i) Nasopharyngeal flocked swab (NPS) method

A modified procedure to that of Chan et al and described in the SSC protocol no. 1527 will be adopted (Chan et al. 2008). This is a simple procedure making use of a double-headed nasopharyngeal swab whose the tip diameter is less than 3mm. Briefly, the distance between the participant's nares and earlobe is measured to estimate the length of insertion. The swab is then gently inserted up the nostril towards the pharynx for the measured distance. The swab is rotated 3 times, to obtain epithelial cells and surface colonising bacteria and held in place for 5 seconds to ensure maximum absorbency. The swab is then withdrawn gently and put in an appropriate transport medium. One tip of the swab will be placed in skim-milk

tryptone glucose glycerol (STGG) transport media and the other into viral transport media and processed with minimal delay (<6 hours) at the KEMRI CGMR-Coast laboratory.

ii) Oral fluid (OF) collection procedure

Oral fluid will be collected using a sponge swab (Oracol, Malvern Medical Developments, Worcester, UK), consisting of a cylinder of expanded polystyrene foam attached to a plastic stick and is used like a toothbrush (Nokes et al. 2001; Okiro et al. 2008). The swab is brushed along the gums and mouth for 60 seconds and the device is then inserted into a plastic tube, stoppered, stored on ice and returned to the laboratory upon return from the field. In the laboratory 1ml of preservative buffer (20% foetal calf serum and 0.2% sodium azide in PBS) is added to the sample. The oral fluid is then squeezed out of the foam and clarified by centrifugation and the supernatant stored in a -70° C freezer for processing at a later stage.

Laboratory

All specimens will be immediately stored in a cool box (with ice packs) before being transported to the RSV lab for storage in 4^0C fridge. For longer-term storage the samples will be kept at -70°C.

i) Molecular methods for respiratory viruses

Nasopharyngeal swabs (in viral transport medium) will be screened for RSV A and B, hMPV, PIV3 and rhinovirus and 8 other respiratory viruses using a multiplex (MPX) PCR real time assay system (Gunson et al. 2005). Real time RT-PCR will enable quantification of viral load in virus positive household members (van Elden et al. 2003). G-gene genotyping, selective sequencing, and recombinant expression, will be undertaken for RSV variant molecular and immunological characterisation (Scott et al. 2004; Scott et al. 2006; Scott et al. 2007). Real time PCR methods will also be used for the detection of viral antigens in OF samples. All assays are, or will shortly be, established in the Kilifi laboratory.

ii) RSV Antibody assays (OF)

OF will be tested for RSV specific IgG by optimised indirect ELISA (Okiro et al , 2008). Interpretation of changes in specific antibody concentration and the definition of re-infection will be made based on the antibody profiles as described elsewhere (Okiro et al. 2008)

iii) Nasopharyngeal bacterial carriage tests

The sample placed in skim-milk tryptone glucose glycerol (STGG) transport media will be inoculated onto appropriate culture media, before storing the residual sample in STGG in a freezer at -80°C. Bacteria will be identified from culture media using standard microbiological/molecular methods. Pneumococci will be identified from gentamicin-blood agar by Optochin susceptibility testing and Quellung serotyping. Haemophilus influenzae will be identified from bacitracin-chocolate agar by X and V factor dependence and slide agglutination serotyping. If necessary, PCR of bacterial isolates will be utilized to confirm identity (e.g., to distinguish non-typeable Haemophilus influenzae from Haemophilus haemolyticus, or to serotype pneumococci (O'Brien and Nohynek 2003; Antonio et al. 2009).

Quality assurance and control

Standard procedures for quality control of assays and apparatus are operating the laboratory and will be monitored throughout. All laboratory staff are trained in Good Laboratory Practice.

Sample storage and exportation

At the end of the study, residues of clinical specimens (nasal samples, oral fluids) and viral isolates will be archived indefinitely in KEMRI-Wellcome Trust research laboratories for future use in the study of RSV or other respiratory pathogens. In addition to assay quality assurance and molecular studies on RSV described above for the present study, it is envisaged that future studies may be undertaken involving laboratory techniques not supported in Kilifi and requiring export to collaborators at the University of Warwick, and Health Protection Agency, UK. A separate proposal will be submitted prior to conducting any future studies on the archived samples either in Kilifi or elsewhere (University of Warwick or Health Protection agency, UK) with relevant documentations. Where reasonably possible, a Kenyan laboratory technician or research scientist will assist or lead in any work being undertaken abroad to strengthen capacity in Kilifi and Kenya.

i) Data Management and analysis

Data Storage

Clinical data will be entered directly onto the Kilifi Integrated Data Management System (KIDMS) existing at the Centre for Geographic Medicine Research Coast, Kilifi. All databases are backed up to a local alternative hard-drive on a weekly basis and to a remote hard-drive at the Wellcome Trust Research Laboratories in Nairobi on a monthly basis. Laboratory data will be stored separately on Filemaker and Excel files and will be backed up on a separate hard drive and CD. Extraction of data for analysis will be via password-protected access to the study investigators under the control of the PI and computer services manager. Laboratory and clinical data will be merged and analysed using STATA v9 (StatCorp, US).

Data analysis

Defining of transmission data and calculation of simple statistics

Using the RSV infection and G-gene sequencing data 'temporal' links of RSV transmission based on established methods for determining possible successive cases will be defined (Fine 2003). This will be repeated for the other viruses. For RSV infection the following estimates will be calculated; (i) the proportion of households at the end of the study with at least one infected occupant, (ii) the proportion of infants, siblings, parents who were infected, (iii) the proportion of primary RSV infections in infants that are from within and from without the immediate household, (iv) for within household transmissions, the proportion of the infants acquiring RSV infection from parents and/or elder siblings. The above estimates will be stratified on the various characteristics such as household size and viral load where numbers permit. Crude attack rates and secondary attack rates of RSV (also for other viruses) within the household stratified by age group and on basis of the initial source of infection (infants,

siblings and parents (mother)) will also be computed. The above calculations will be repeated for the other respiratory viruses.

Estimation of incidence rate of infection and recovery of RSV and other respiratory viruses

Poisson regression methods will be used to estimate the incidence of RSV (and repeated for other viruses) infection (using person-days of observation as denominator) estimating relative rates in (a) different age groups (b) between households of different size (c) previously infected or uninfected groups. The analysis will adjust for age and presumed past history of infection. Similar analysis will be utilized to estimate the rate of loss of shedding (or average duration of shedding) in relation to age, past exposure and other demographic factors. Comparison of our estimates of incidence rates of infection and recovery with those from our previous birth cohort (Nokes et al. 2004; Nokes et al. 2008) and household study (Okiro 2007) will also be made (the original datasets are available for this analysis).

Household reproduction number, R_H and transmission probability

We define a household reproduction number (R_H) as the average number of secondary cases arising from one case throughout the infectious period following introduction into a household. Stratifying the household into two groups (infant and older (siblings and adults)) a reproduction potential matrix will be generated e.g. R_{Hio} is the number of secondary cases among the older (o) ages arising from an introduced infant (i) case. The R_H will inform on the household transmission potential, where an average $R_H<1$ would indicate little potential for household spread.

We can estimate the probability of transmission per separate contact between infant and older individual, from the household reproductive potential. If c_{io} is the number of contacts made between the infant and older individuals per day, throughout the infant infectious period, D_i and p_{io} is the risk of acquiring infection per contact event, then the infant household reproduction potential can be defined as

$$R_{Hi} = p_{io}c_{io}D_i \qquad (1),$$

Hence, with independent knowledge of the average duration of infection, the household reproduction potential and contact rates the transmission probability per contact can be estimated. This can be modified to identify the transmission probability per contact with a different individual (rather than per separate contact). The same can be estimated for the introduction of an infected older individual, R_{Ho} where

$$R_{Ho} = p_{oo}c_{oo}D_o + p_{oi}c_{oi}D_o \qquad (2),$$

bearing in mind that older individuals can infect the infant and other older individuals. We intend to estimate the rates of contact (total or new) between individuals through parallel contact pattern studies.

Data on other viruses will allow us to determine more generally the importance of certain groups in household introduction and transmission, and the relationship between contact pattern data and transmission of respiratory viruses.

Evaluation of the diagnostic performance of OF in detecting viral infections using PCR methods

Paired OF and NPS samples will be screened using RT-PCR for RSV and other respiratory viruses. The diagnostic performance of OF PCR evaluated against NPS PCR (used as the gold standard). Sensitivity, specificity, positive predictive value and negative predictive value of the OF PCR will be compared using McNemar's chi-square test. The effect of age and disease severity on test performance will also be checked through stratification.

Evaluation of the association between RSV infection and acquisition of pneumococcus or haemophilus influenza

Using Poisson regression methods we will compare the rate of bacterial (pneumoccus or Haemophilus influenza) nasopharyngeal colonisation following an RSV infection with those who have not had RSV infection. Similar analysis will be repeated for the other viral infections where the numbers permit.

j) Time Frame/Duration of the Project

- Proposal development and Ethical approval: March to July 2009
- Study preparations – Field worker recruitment and training: August to September 2009
- Mapping exercise and initial piloting phase: October 2009
- Data collection: November 2009 to April 2010 (depending on the duration of the RSV epidemic)
- Laboratory screening of samples: November 2009 to July 2010
- Data analysis and final report writing: August 2010 to December 2011

k) Ethical Consideration

Risks.

No procedures used in this study represent a significant risk to the participants. The nasal sampling may result to mild discomfort which will be explained to each participant prior to collection. Suitably trained staff will undertake all procedures, which will be carefully explained in advance to the participants and their children (as appropriate). The frequency of home visits will be discussed with the household head and all individual members so that a convenient schedule is agreed upon. We aim to recruit staff from the study area who would understand the local culture.

Benefits

There are no direct benefits accruing from the frequent nasal/oral sampling and testing for there is no established treatment for viral respiratory infections. However, participants will benefit from close monitoring for any illness by a nurse during the home visits and will be referred to the health centre promptly for treatment where necessary. Costs of treatment of acute illnesses incurred in the nearby government health facility will be paid for plus costs of referral. We will address the issue of how to compensate the participants for their time through a process of consultation with local community leaders and representatives, facilitated by the centre's Community Liaison Group. We will seek to provide this compensation in kind where this is considered appropriate.

Community benefit

Benefits to the community are general and the effects will not be immediate. We aim to generate data that will inform the development of vaccines, and are useful to the public health policy makers in optimising the use of a future RSV vaccine.

Consent

Consent will be obtained from the household head and all members of the household only after adequate explanation of the study (see Appendix G). The study will be explained to those aged less than eighteen years (who are able to understand the study) and their verbal assent obtained before a signed informed consent is obtained from their parents or guardians. The individual informed consent forms will be translated into Giriama and Kiswahili and administered in the appropriate language, to all the study participants by trained field workers – conversant with the local languages. Specific consent will be sought for storage and later use of samples, and for exportation of samples. It will be explained that use of archived samples and associated data, and exportation, for new studies, will only be possible after scrutiny and approval from the independent National Research Ethics Committee.

Confidentiality

Data will be stored on password-protected database, accessible only to study investigators. Data will be archived with the database administrator.

Community engagement strategy

A community engagement advisory group (CEAG) spearheaded by the centre's community liaison department will oversee community sensitisation and provide support during the study implementation (Marsh et al. 2008). Multiple approaches including holding (i) consultative meetings with the community leaders and representatives, (ii) public meetings and (iii) meetings with the household heads are some of the suggested community engagement strategies. It is essential to establish communication channels and foster a favourable environment in the community for this intensive study to be successful.

Training/support for those involved in community engagement

Fieldworkers and clinical team involved in the household recruitment and collection of specimens are all familiar with the goals of KEMRI and the consent process. Any additional staff will be trained on communication process and consenting procedures. We will work in conjunction with the KEMRI community liaison group, and the CEAG will provide support to our community activities.

Feedback of information

The individual test results of the nasal and oral samples will not be conveyed to the participants for they do not affect their health care. This will be clearly explained during the consenting process. A feedback meeting will be organised at the end of the study to share the major study experiences and results. The study findings will be published in peer reviewed journals.

Animal Subjects

Not applicable

l) Expected Application of the Results

The findings will inform mathematical models on RSV transmission dynamics and in evaluation of population level impact of various vaccination strategies. This later analyses will be useful in informing the public health policy makers in ways to optimise the use of a RSV vaccine in future.

m) Budget

Item	KSH	GBP
1. Personnel – salaries and other benefits	2,406,162	20,923
2. Patient costs, travel, food and/or supplies	50,000	435
3. Equipment- motorcycles and bicycles	976,772	8,494
4. Supplies and laboratory tests	17,624,040	153,253
5.. Transportation, vehicle repairs, insurance etc	249,716	2,171
6. Operating expenses - postage, printing etc	82,000	713
7. Animals acquisition, food, cages etc	N/A	N/A
8. Consultancy fees	N/A	N/A
9. Contingency (15% of above total)	3,203,534	27,857
10. Institutional administrative overheads*		
Total	24,560,424	213,846

* The project forms part of the KEMRI/Wellcome Trust Collaborative Programme and will attract no overheads.

Justification of the Budget

Personnel Salaries and Benefits - this includes cost of hiring project coordinator and study fieldworkers to help in the study implementation for a period of one year.

Patient costs, travel, and food and/supplies – this includes medical expenses in the outpatient clinic, fare refunds and cost of the possible compensation packages for the study participants.

Equipments – This includes the cost of purchasing new motorcycles, bicycles, desktop (for data entry) and laptop (for the ARO) personal computers. Mobile phones are also included.

Clinical supplies and laboratory tests – this is the estimated cost of study supplies (nasal swabs, rubber bulbs, gloves etc plus the cost of the 13+ multiplex virus screen, real time RT-PCR, oral fluid ELISA and bacterial work.

Transportation, vehicle/motorcycle repairs, insurance - includes the estimated cost of fuel, servicing, insurance and riding gears of the motorcycles, bicycles, and vehicles used in the participant follow-up.

Operating costs - these includes all costs of photocopying, stationery, printing and all communication costs. Costs incurred during the community engagement and sensitisations are also included.

n) Attachments:

- Role of investigators
- Information and consent form for the Household study (included in this thesis as a stand alone appendix

o) Role of investigators

i) Patrick Munywoki (Assistant research Officer); the principal investigator with the overall responsibility for all aspects of the work.

ii) Dr James Nokes (Infectious disease epidemiologist); responsible for the project conception and involved in protocol design, data collection and analysis and final report writing.

iii) Charles Nyaigoti (Assistant Research officer), Anne Bett (senior Lab. Technician): Laboratory staff involved in molecular diagnosis of respiratory viruses. Charles Nyaigoti will be particularly involved in establishing of real time RT-PCR methods.

iv) MJ Ngama (Clincial Officer) and Dorothy Koech(Graduate Nurse) are responsible for study coordination including field worker training (consenting process and study procedures), liaising with community leaders and offer clinical care to the participants. Contributors to study implementation and report writing.

v) Dr Laura Hammit (Paediatrician): Contributor to study design, implementation, and analysis and report writing. Provide clinical support.

vi) Dr Pat Cane (molecular biologist): will provide molecular biology support in the UK, including quality control. Undertake G gene genotyping, sequencing, and GST fusion protein expression.

vii) Prof Graham Medley (Infectious disease epidemiologist): Contributor to study design, epidemiological analysis and report writing. Based in the University of Warwick, UK.

Appendix B. Estimated proportion shedding at different sampling intervals

The estimated proportion of individuals who remain shedding virus after t days post onset of shedding (St), where the (constant) rate of recovery is r, s(t) is given by:

$s(t) = \exp(-rt)$.

Assuming the average rate of recovery (cessation of shedding) is the reciprocal of the mean duration of shedding (i.e. between 3.5 and 9 days), and that shedding onset occurs on average mid-way between the sampling interval on average, then we generate the following relationship in Figure 2 between shedding duration (i.e. r=1/3.5, 1/7 and 1/9 day^{-1}) and proportion shedding (S(t)), for intervals of 3.5 days (i.e. t=1.75 days) and 7 days (i.e. t=3.5 days)

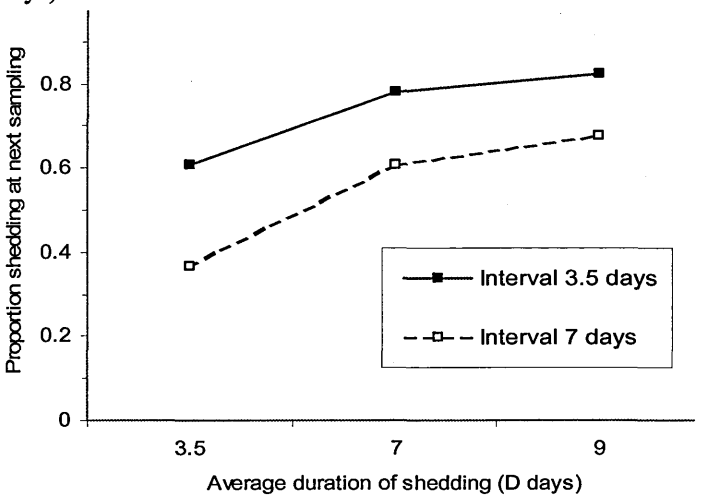


Figure: Proportion of shedding at the next sampling by average duration of shedding by sampling interval

The proportion of individuals predicted to remain shedding, and thus detectable, will range from 61%-82% (for 3.5-9 days duration) for a 3.5-day sampling interval (Figure). The comparable range is 37%-68% for a 7-day interval. Given the need to detect infection in mild cases and in older children and adults with likely lower range of shedding duration, sampling twice weekly was indicated.

Appendix C. Household study flyer:

A KEMRI study on how a germ (RSV) causing pneumonia passes between people in families

a) What is KEMRI?

KEMRI is a Government organization under the Ministry of Health that carries out health research to learn more about diseases that affect children and adults in Kenya. Scientists working at the programme come from all over Kenya, East Africa, and other countries worldwide. Research results are used in Kenya and throughout the world to improve health. All studies conducted at KEMRI Kilifi have been approved in advance by expert committees in Kilifi and Nairobi who check the research question is important, the study will be conducted well and that the safety and rights of participants are protected.

b) What is this study about?

This study is about finding new ways of protecting children from one type of pneumonia. Pneumonia is one of the most common serious illnesses in Kenya and other parts of Africa, causing cough, fever and difficulty in breathing. It is usually caused by different types of germs that enter the body through the nose. Pneumonia germs pass from one person to another, particularly when they cough or sneeze. Respiratory syncytial virus (RSV) is one type of germ that can cause pneumonia in young children. However, RSV does not always cause illness. Commonly people (especially older children and adults) have this germ in their noses and throats without any symptoms at all. However, when they pass it on to someone else, that person may develop symptoms. This study will also find out more about why it sometimes causes illness and sometimes does not. RSV occurs mainly during the season between November and May. Every year about 1 in every 100 young children in Kilifi District are admitted to the hospital with pneumonia caused by RSV. At the moment, RSV pneumonia has no specific treatment so it is very important to find ways of protecting children from this infection. A vaccine for RSV is under development at the moment. In order to know the best way of using such a vaccine, we need to understand how the virus spreads from person to person, particularly in families. This study aims to find out the way that RSV spreads between members of 50 families living in this area who volunteer to take part.

c) What does the study involve for those who are in it?

This study will be conducted in Matsangoni location. It will start around October 2009 and continue for 4 or 5 months. We would like all family members and other permanent residents from 50 households to participate. In order to participate, families must have a child under one year and at least one older child. It is also important that they will be available to take part from the beginning to the end of the study. For families that agree to participate, a field worker will visit the home twice a week for 4 to 5 months to collect samples from the noses or mouth of all household members. We will make sure that samples are taken at a time that causes as little inconvenience to families as possible. Families will also be visited weekly by a nurse to advise on health issues. KEMRI will provide some types of treatment for family members who become unwell during this study. All details of the study will be explained to

families before they decide if they would like to participate. By participating in this study, families will be helping to find a vaccine for RSV pneumonia that could benefit many children in Kenya in the future.

Appendix D. Frequently asked questions about RSV

This leaflet contains salient information about RSV infection and disease.

a) Introduction

Respiratory syncytial virus, or RSV, is a respiratory virus that infects the lungs and breathing passages. Most otherwise healthy people recover from RSV infection in 1 to 2 weeks. However, infection can be severe in some people, such as infants, young children, and older adults. In Kilifi, we have found that RSV is the most common cause of severe respiratory illness in children under one year of age.

b) What is RSV?

RSV causes upper respiratory infections (such as colds) and lower respiratory tract infections (such as bronchiolitis and pneumonia). In children under 1 year of age, RSV is the most important cause of bronchiolitis, an inflammation of the small airways in the lung.

Almost all children will have had an RSV infection by their second birthday. When infants and children are exposed to RSV for the first time, 25% to 40% of them have signs or symptoms of bronchiolitis or pneumonia, and 0.5% to 2% will require hospitalization. Most children hospitalized for RSV infection are under 6 months of age.

c) Is RSV an important cause of respiratory illnesses in Kenya?

RSV is a very common cause of respiratory illness in Kenya and in Kilifi in particular. Each year, one in every 100 children in their first year of life is hospitalized with severe RSV disease in Kilifi District Hospital.

d) How is RSV spread?

RSV can be spread when droplets containing the virus are sneezed or coughed into the air by an infected person. Such droplets can linger briefly in the air, and if someone breathe in the particles or the particles contact their nose, mouth, or eye, they can become infected.
Infection can also result from direct and indirect contact with nasal or oral secretions from infected persons. Direct contact with the virus can occur, for example, by kissing the face of a child with RSV. Indirect contact can occur if the virus gets on an environmental surface, such as a doorknob, that is then touched by other people. Direct and indirect transmissions of virus usually occur when people touch an infectious secretion and then rub their eyes or nose. RSV can survive on hard surfaces such as tables and doorknobs for many hours. RSV typically lives on soft surfaces such as tissues and hands for shorter amounts of time.
People infected with RSV are usually contagious for 3 to 8 days. However, some infants and people with weakened immune systems can be contagious for as long as 4 weeks. Once RSV is introduced in the household is it can be rapidly transmitted to other members of the family, often infecting about 50% of other household members.

e) What are the symptoms of disease caused by RSV?

Symptoms of RSV infection are similar to other respiratory infections. A person with an RSV infection might cough, sneeze, and have a runny nose, fever, and decrease in appetite. Wheezing may also occur. In very young infants, irritability, decreased activity, and breathing difficulties may be the only symptoms of infection. Most otherwise healthy infants

infected with RSV do not need to be hospitalized. In most cases, even among those who need to be hospitalized, hospitalization usually last a few days, and recovery from illness usually occurs in about 1 to 2 weeks.

f) Who is at risk for severe illness?
Premature infants, children less than 6months, children less than 2 years of age with congenital heart or chronic lung disease, and children with compromised (weakened) immune systems due to a medical condition or medical treatment are at highest risk for severe disease. Adults with compromised immune systems and those 65 and older are also at increased risk of severe disease.

g) When is the risk for infection the greatest?
RSV infections generally occur in Kilifi, Kenya from November to April. However, the timing of the season may differ among locations and from year to year.

h) How can I provide care to someone with RSV?
There is no specific treatment for RSV infection. However, there are simple ways to help relieve some of the typical symptoms. Your doctor can give advice on how to make people with RSV infection more comfortable and assess whether hospitalization is needed.

i) How is the infection prevented?
Researchers are working to develop RSV vaccines, but none is available yet. However, there are steps that can be taken to help prevent the spread of RSV. Specifically, people who have cold-like symptoms should:

- Cover their mouth when they cough and sneeze
- Wash their hands frequently and correctly (with soap and water for 15–20 seconds)
- Avoid sharing their cups and eating utensils with others
- Refrain from kissing others

In addition, cleaning contaminated surfaces (such as doorknobs) may help stop the spread of RSV.

Appendix E. Field worker training timetable

Ref	Content	Objectives	Who	Required	Preparations to be made before - requirements	How long/day
Module 1	Opening the workshop	To create a positive environment for participatory learning	CLG	Newsprint Marking pens Projector Name tags Pre-test questionnaire	Objectives of the training (transparency or newsprint) Training programme (Handouts)	1.5 to 3 hours
Module 2	What is KEMRI?	To strengthen participants' understanding of the research Centre, its aims and core values, its linkages and its relationships with stakeholders	CLG	Newsprint 4 A4 manila papers Marking pens Projector Video on Community Perspectives of KEMRI Name tags Pre-test questionnaire	Objectives of the training Newsprint or slide on KEMRI's main aims	3.5 hours
Module 3	More about Health Research: Specific aspects of research	Introduction to basic principles of health related research, based on an		Newsprint Markers	Prepared "Research or not	4.5 hours

	activities, (focuses on aspects commonly misunderstood)	awareness of the participants' current knowledge, understanding and experience	CLG	Handouts	research" scenarios Slides of different research activities Review process Role play instructions	
Module 4	Protecting the participants rights in research	To strengthen understanding of the history and development of ethical principles/guidelines for research involving people	CLG	Newsprint Marker pens Role play guidelines Handouts	Definitions of principles of research ethics Benefits vs Risks in research Case studies	3hours
Module 5	Communication Skills	To strengthen understanding of the definition and components of communication	CLG	Newsprint Marking pens Pre-test questionnaire Handouts	Cases for Role play	3hours
Module 6	Answering Questions and Solving Problems	Strengthening awareness of issues involved in effective communication, including skills for answering questions and solving problems	CLG	Newsprint Marking pens Role Play guide Handouts	Cases for Role play	1.5 hours
Module 7	Introduction to Household Study	A brief introduction to the study.	PI	Newsprint Marking pens Study manual	FW manual	1.5 hours
Module	Household Consent Training	To develop communication skills for		Newsprint	FW manual	3 hours

8		handling problem situations in normal workplace	PI	Marking pens Role Play guide FAQs Handouts	Role play	
Module 9	The Home Visit: Sample collection and illness assessment	Step by step guide on what is required once the field worker arrives at the participants home	PI	Newsprint Marking pens Role Play guide SOPs Handouts	FW manual Sample collection devices	All day

Key: PI, principal investigator; CLG, community liaison group; FW, field worker; SOPs, standard operating procedures; FAQs, frequently asked questions

Appendix F. The field workers' manual

Household transmission of respiratory syncytial virus (RSV): Who acquires infection from whom

Compiled by: Patrick Munywoki, Dorothy Koech, and D James Nokes;

First draft on 20th July 2009; last updated on 10th January 2010

I. Introduction

KEMRI is a Government organization under the Ministry of Public Health and Sanitation that carries out health research to learn more about diseases that affect children and adults in Kenya. Scientists working at the programme come from all over Kenya, East Africa, and other countries worldwide. Research results are used in Kenya and throughout the world to improve health. All studies conducted at KEMRI, Kilifi, have been approved in advance by expert committees in Kilifi and Nairobi who check the research question is important, the study will be conducted well and that the safety and rights of participants are protected.
The current study is organised by KEMRI-Wellcome Trust Research Programme in collaboration with University of Warwick, England. We are investigating an infection caused by a germ (virus) that is known to be a major cause of chest (respiratory) infections through out the world. The virus is called Respiratory Syncytial Virus or RSV.
RSV is considered the major pathogen causing severe lower respiratory infections among infants and young children worldwide. It is the single most important cause of pneumonia and bronchiolitis in infants and young children less than one year of age. For infants who are infected in their first year of life, the symptoms can frequently be severe enough for the parent to feel the need to take the child to hospital. The virus is unusual in that it can repeatedly re-infect individuals throughout life though subsequent infections are associated with decreasing severity. Because of the severity of RSV disease, its treatment and prevention are important public health concerns.
Studies from Kilifi District show that around 1 in every 100 infants are admitted to hospital with RSV associated severe pneumonia each year and there is a clear seasonal pattern on the occurrence of RSV epidemics every year. RSV infections usually start in November and peaks in January-to-February lasting about 4 to 5 months.
We have currently an on-going study in children wards in Kilifi District Hospital (KDH). Briefly, this is a long-term surveillance of Respiratory Syncytial Virus (RSV) in children admitted with lower respiratory tract infections (LRTI) to KDH which aims at:

a) Determining seasonal and longer-term periodicity of RSV infection,
b) Obtaining an accurate estimate of RSV associated mortality in acute respiratory infection (ARI) admissions,
c) Monitoring group and genotype prevalence, and relating variant type to severity,
d) Creating a panel of RSV isolates, with associated acute and convalescent sera, and stratified by patient HIV status, suitable for functional strain-specific immunity assays,
e) Comparing RSV disease in HIV infected and uninfected children,
f) Investigating other viral causes of ARI admissions in children less than 2 months old.

A vaccine is not yet available to prevent RSV infection but is under development. The potential usefulness of such a vaccine and information on how best it can be implemented requires a better understanding of the spread of the virus within the community. The family or household unit is an important element in the spread of infections, transmitted by close contact, in the community. Better understanding of household transmission may inform on the merits of targeted control strategies. Data on within-family spread of RSV, particularly on who infects the infant, are scarce. Close monitoring of infections within families will provide this vital information hence need for household-based studies

II. The household study

There are a number of studies planned to investigate the spread of RSV within the community. A brief summary of household study, which has since been approved, is presented in this manual. The study seeks to understand the spread of RSV within the community particularly in households generating data that can help in assessing who infects the young children and particularly the infants. The study will be conducted during the forthcoming season of RSV likely to start from November 2009 to May 2010.

III. The Household study: An overview

a) Study Area

The study will be conducted in Matsangoni Location, Kilifi district, mainly within Uyombo and Mtasangoni sublocation but may spill over to the neighbouring sublocation near Matsangoni dispensary i.e. Mkongani (these may be revised). The community strategy (every 20 households are under surveillance of a Community Health Worker) spearheaded by the Ministry of health has been successfully implemented in making it easier to identify the eligible households and carry out community sensitisation.

b) Aims of Study

The focus of the study will be to monitor RSV infections within a group of households with an infant and at least one older child (ren) during a season of high transmission of RSV(4 to 5 months period starting in November 2009). The study aims to **to understand who acquires infection from whom** (WAIFW) by answering the following questions:

i) Who brings infection into the household?
ii) Who gets initially infected in the household?
iii) Who infects whom?
iv) What proportion of the household members become infected, and, in particular,
v) Who infects the infant?

In particular, we are interested in assessing whether it is school-age children who are source of infection for their younger housemates (usually siblings) who are at greater risk of subsequent disease. Nasal swabs and oral fluid samples will be collected from each member of the household regardless of their symptoms.

c) Study design

This is a household-based cohort study involving about 50 households with an average of 7 individuals per household. Recruitment procedures will include community sensitization, identification of households, consenting of the participants and collection of demographic data.

The participating households will be followed up during the RSV season (the start identified by an increase in RSV-associated admissions in KDH, and probably running for about five months). We will make **twice a week home visits** in which each household member will have the samples collected (oral and nasal). These specimen will help us determine whether RSV is present and will also determine the strain of the virus if they are currently infected. The household members will further be asked about their health status and history of any respiratory symptoms. Any sick individuals will be referred to the Matsangoni dispensary where they will receive treatment free of charge.
Once again, the main aim of this is to identify every RSV infection within the households to be able to determine who acquires infection from whom and in particular, who infects the infant.

IV. The roles of a field worker

a) General
As a field worker (FW), it is your job to ensure the smooth running and success of the study. It is crucial that you fully understand what the study is all about so that you will be able to handle any issues that arise during your work. **You are a very important person for the success of the program and your dedication and commitment will determine the outcome of the study**. You will therefore be expected to present yourself well to the community so as to be respected and accepted within the area. As you visit each home, be courteous and make the participants comfortable.

i) Always introduce yourself.
ii) Always wear your KEMRI T-shirt (if you don't have one we will provide).
iii) Always wear your KEMRI ID.

b) Skills you will need for this study
The various specific skills required for this study are listed and details are provided elsewhere in the document:

i) Able to deliver accurate information to the participants or parent/caretaker (if child) and obtain informed consent.
ii) Conduct simple questionnaires
iii) Taking of vital signs, temperature taking, respiratory rate taking, elicit history of respiratory illness, in particular accurately assess breathing rate.
iv) Collect nasal specimen by use of the nasal flocked swab
v) Collect oral-fluid specimen by use of the ORACOL device.
vi) Be able to correctly deal with questions arising from either the study and non-study participants in the community (and refer appropriately)
vii) Ability to correctly deal with issues arising that may be associated to the commitment of the households.

V. Informed consent

Consent for participation of household members into the study will be obtained at the home. The consent form includes considerable detail about the study, the level of involvement of the household in the study, including the advantages and disadvantages of participating.

You may be required to obtain informed consent, and the family/guardian/parent may ask you questions about it then or at a later stage, when you make home visits. You therefore need to understand the information sheets and consent forms, how they are delivered, and to be able to answer questions from the household members on the issues given above. All participants receive a copy of the signed consent form and the study information sheet. Informed consent can only be taken from persons above 18 years. In the case of children the respective parents or guardians will have to provide consent. When providing consent, the following points should be emphasized:

i) Participation in the study is voluntary and anyone is free to. It is necessary if the study is to be effective that the household remains in the study for the entire study period.
ii) Fieldworkers will visit the homes of the participants twice a week during the RSV season (which might be half the year) for the purpose of collection of samples as well as identifying illness related symptoms of each of the household members.
iii) The study will run for about 4-5 months and will then stop.
iv) We will need to collect a nasal sample from each member of the household twice a week.
v) An oral sample will be collected from each of the household members once a week.
vi) Field workers are not medically trained. You should not be asked for advice on the child's other household members' medical conditions neither will you carry any drugs to the field. If any of the participants is ill during the period of study refer them to the Matsangoni dispensary where we will provide definitive treatment free of charge.

VI. Home visits by field worker

This is the main component of the household study. There are certain actions that need to be taken during each of the home visits made by the field worker.

a) Before you depart for the households you should ensure the following:

- Obtain list of households to be visited on that including details of how to locate them.
- Fill out date of the visit at top of the surveillance form and date of next appointment (3 -4 days)
- Confirm whether you will be collecting both oral and nasal samples or only nasal samples for that particular visit.
- Check to take adequate materials/supplies for nasal swabs and oral fluid sample collection (See Appendix for list of required materials.)

b) At the household

- Remember to introduce yourself to the family/mother/head of household and identify yourself as from KEMRI for the household study.
- Remember to always be polite and to respect the wishes of the family.
- Enter participants **study ID number** and **NAMEs/SEX** into the illness assessment form. Confirm the data to ensure you are accurate.
- Conduct an illness assessment for each of the household members just before you collect the samples.
- Inform the participants of what you intend to do and request for their permission before proceeding.
- Collect the samples from each participant in a secluded area away from the other household members.

- Fill in the details of sample collection in the form
- Thank the participant for their cooperation and ask if they have any concerns or questions.
- Plan for the date of next visit (including time), which should be within 3-4 days and find out whether they will be available. If not you may have to schedule the visit at a different time (After discussion with the participant or parent if the participant is a child).

c) Sample collection

The sample will be collected following the respective approved standard operating procedures (SOP). The SOPs for Nasopharyngeal flocked swab (NPS) and oral fluid (OF) collection procedures are attached.

d) Household Illness Assessment form

This is a form that a field worker will be required to fill during each of the visits to the households. It serves the purpose of documenting whether specimens have been collected as well as an illness assessment of each of the household members. Refer to appended form and notes for further details.

The key symptoms you should be able to identify:

- Difficulty in breathing-whose signs are nasal flaring, use of accessory muscles, e.g. neck muscles and chest in drawing (sucking in of lower bony chest wall on inspiration) and wheezing [This is a high-pitched whistling sound made when air flows through narrowed airways in the lungs, usually when people breathing out].
- Cough (less than 30 days);
- Fever (hot body or axillary ≥37.5C);
- Nasal congestion/ discharge;
- Fast breathing; Breathing rate is ≥ 50 AND age < 12 months OR breathing rate is ≥ 40 AND age ≥ 12 months (WHO guidelines).

If a child appears to have **any cough or difficulty in breathing or fast breathing,** refer to Matsangoni dispensary by completing the RSV referral note in RSV field book. Inform the participant he or she will receive fare reimbursement at the dispensary. Emphasize that the participant or child should go there the same day.

Remind them to come to the clinic with their study ID cards.

The nasal specimen should be transported back to the lab **AS SOON AS POSSIBLE.**

VII. Accuracy in data collection and storage of samples

Accuracy in data collection is a top priority. The following are some simple guidelines:

Label all bottles and tubes for saliva and nasal swab collection just before collection of sample– it is easy to forget to label materials in the noise and confusion of sample taking. All samples should be labelled with the individual's name, serial number and the date and time of collection of the sample. **You are discouraged from labelling all the samples tubes together before the collection as this may lead to placing the samples in the wrong tubes.**

- Make every effort to return collected specimens to the research unit as quickly as possible, maintaining cold conditions as required.

- Complete forms and questionnaires carefully. Complete all sections required. Do not be hurried. Do not make assumptions about answers elicited - if unsure, return to the question and ask again.
- Be methodical in your actions. Make sure you are clear about exactly what you intend to do before arriving at the household.

On return to the dispensary

- Go through the forms to ensure that there are no blatant mistakes and hand in the forms to the field coordinator.
- Relate all the problems encountered in the field to the study coordinator and also note them down for discussion.
- Enter study numbers into logbook in the dispensary of individuals from whom nasal swabs and oral fluid samples have been collected.

VIII. Procedures

a) Temperature Recording

All temperatures will be taken from the armpit (axillary temperatures). Do not round off the figures on the thermometer but record it as it is on the forms.

b) Recording breathing rates

Defined as the number of breaths per minute. Obtained by observing the frequency of inspiratory phase over 30 seconds then multiply by 2.

c) Collecting of specimens

Nasopharyngeal swabs (NPS) (Refer to SOP for detailed explanation of the procedure)

The NPS is a simple procedure making use of a commercially available flocked swabs. Briefly, the distance between the participant's nares and earlobe is measured to estimate the length of insertion. The swab is then gently inserted up the nostril towards the pharynx for the measured distance. The swab is rotated 3 times, to obtain epithelial cells and surface colonising bacteria and held in place for 5 seconds to ensure maximum absorbency. The swab is then withdrawn gently and put in an appropriate transport medium.

Oral fluid (OF) samples (Refer to SOP for detailed explanation of the procedure)

The OF will be collected using a sponge swab, consisting of a cylinder of expanded polystyrene foam attached to a plastic stick and is used like a toothbrush. The swab is brushed along the gums and mouth for 60 seconds and the device is then inserted into a plastic tube, stoppered, stored on icebox.

d) Storage of specimens

Immediately after collection, and securing screw cap lid tightly, put the specimen in the icebox. Transport the specimen to the lab together with the filled microbiology request form to the lab as soon as possible.

IMPORTANT: Contamination between containers, gloves, and swabs should be avoided. Very sensitive techniques will be used to detect the virus and any contamination will be detectable. So, DO dispose off gloves, or store separately used ones, after each procedure, make sure lids of vials are in firmly, DO NOT reuse storage tubes.

IX. Anticipated challenges

Fieldwork rarely goes as smoothly as is planned and if we can anticipate some problems and know how to go round them earlier in the study, it will make work easy for us. As the study continues, more problems are sure to arise. We will discuss these problems as a group once a week and decide ways in which to handle them. Do not give random explanations when accosted with a question that is difficult to answer. If you do not know the answer to an issue that is raised say so, note the comment on a piece of paper and we will discuss it to allow for uniformity of our answers. 'White lies' may lead to a lot of complications in the future and should be avoided since we intend to stay in the study area for a long time. Some of the problems anticipated include:

a) Absent participants

If no any of the participants is present in the house on arrival, check at the neighbours if the person is in a nearby shop or has gone to fetch water and if all these are close by, wait for the person. If the person is still not available, follow up later in the day. Meanwhile collect samples from the household members available. If they have gone to the shamba, find out how far it is and whether it is possible to go there.

b) Very sick people in need of urgent treatment

If you come across a very sick person, whether or not they are a study member, the golden rule is to seek help from the nearest available source. **You are not trained to manage sick individuals and should not take responsibility for sick patients.** In most instances, the first line of help should be the nearest clinic which is staffed with qualified nurses from 8.00 am - 5.00 pm during the working week. If for any reason the clinic is closed, you should advise the patient to seek help at the hospital by the quickest available transport. This will usually be a Matatu or private vehicle. If the Unit vehicle is, soon due to return for pick up, it may be appropriate to help transport the patient with you to Kilifi District Hospital. However, this should only happen under **exceptional circumstances, as passengers are not authorized.**

c) People asking for drugs

Let everyone know that you will not be carrying any drugs with you as you do your work. Those study participants that have any common illnesses will be referred to Matsangoni dispensary where they will receive free medical care and their fare to the place will be reimbursed. Severe cases may further be referred to Kilifi District Hospital. This will only be applicable for acute illnesses.

Participants with other conditions not urgent but needing care and advice seeking on management of diseases

If any of the participants may have illnesses (such as scabies, head lice, minor injuries etc) you should advise them to seek help preferably from a medical professional. You should make it clear that you are not trained to give such advice yourself. You should the participants that they have two main choices:

- **Matsangoni dispensary:** Participants will be encouraged to come for assessment at the matsangoni dispensary during normal opening hours whenever they are ill. We will pay the bus fare and they will not have to pay the normal user fee and will receive routine treatment without charge. If we cannot treat them at the clinic,

(including that the appropriate drugs are not available), we will refer them to the Kilifi District hospital and will also cater for their medical expenses.

- **To seek advice through the normal sources** (usually a shop, herbalist, the nearest private clinic). Some participants may decide to use other preferred sources of health care, which may be more convenient for them. In such cases, it is important to remind them that we will not reimburse for such costs.

d) Anger from the excluded the households in the homestead.

Such households will be reminded that the study requires households that have an infant and an older child. If they are eligible but were not included in the study, inform them that research work had funds for only a particular number of people.

X. Attachments (note: Included as stand alone appendices in this thesis)

- Study Information sheet and consent forms-English
- Study Information sheet and consent forms-Swahili (not presented in this thesis)
- Study Information sheet and consent forms-Giriama (not presented in this thesis)
- Initial Home Visit form
- Home Visit form
- Clinic Visit form
- Nasal Flocked Swab Procedure SOP
- Oral Fluid Sampling Procedure SOP
- Household Risk Survey questionnaire

Appendix G. Household study information sheet and consent forms

I. Household study Information Sheet

a) Study title

Household transmission of respiratory syncytial virus (RSV): Who Acquires Infection From Whom

b) Investigators

Institution	Investigators
KEMRI CGMRC, Kilifi	Principal investigator: Patrick Munywoki Co-investigators: D James Nokes, Laura Hammitt, Mwanajuma Ngama, Dorothy Koech. Ann Bett, Charles Nyaigoti
University of Warwick	D James Nokes, Graham Medley
Health Protection agency, UK	Pat A Cane

c) What is KEMRI and why am I here?

I work for KEMRI. KEMRI is a government organisation that carries out medical research to find better ways of preventing and treating illness in the future for everybody's benefit. One illness KEMRI is currently trying to learn more about is pneumonia caused by 'germs' called respiratory syncytial virus (RSV), which commonly affects children. These germs are spread through close contacts via large nasal droplets from infected persons. A vaccine is not yet available to prevent RSV infection but is under development. In order to know which group to target for vaccination in future when that vaccine becomes available, we need to know how the virus is spread within the community. We aim to closely monitor respiratory infections within members of the household in order to understand who transmits the RSV infection to young children.

To help us learn more we are asking members (about six per household) of 50 households, to allow us to take nasal mucus samples two times per week, and saliva samples once a week for 4-5 months when RSV infections are common. We have selected to ask you since your household has a child aged less than one year and with at least one elder sibling to the child.

d) What will it involve for me/my household if we agree?

If you agree, we will take a sample of mucus from the nose twice a week and saliva once a week (only two home visits per week) for a period of about 4-5 months from all members of the household. The nasal specimen will be collected by inserting a sterile cotton swab in one nostril and twisting 3 times before gently removing it (takes about 10 seconds) – show devices and methods. The saliva specimen will be collected by inserting a sterile cotton swab in one's mouth and rubbing against the gums (as in a toothbrush) for a period of one minute. [Show the device and the method].

During every home visit, we would also like to ask you some simple questions about you/your household related to history of respiratory illness, and take a temperature.

e) Are there any risks or advantages to my household or me if we participate?

When we collect a nasal mucus sample you/your child may experience some mild discomfort but there are no known risks associated with taking this sample. There are also no known risks involved when taking saliva sample from the mouth. Trained and experienced staff will undertake all the procedures. Frequent home visits might cause some inconveniences but considerate schedules based on agreed times will be set.
You/your child will benefit from close monitoring for any illness by a nurse during the home visits and will be referred to the health centre for further consultation and treatment where necessary. Medical expenses for outpatient treatment of acute illnesses will be paid for and where applicable transport costs to the nearest hospital will be met by KEMRI. These benefits will be applicable during the period of participation in this study only.

f) What happens if I refuse to participate?
All participation in research is voluntary. You are free to decide if you /your household can take part. If you agree you can change your mind at any time and withdraw (yourself/your child/your household) from the research. This will not affect you/your child's/your household's care now or in the future.

g) What happens to the samples?
Individual names are removed from all samples and replaced by codes, to ensure that samples can only be linked to the participants by people closely concerned with the research.
Most of the research tests on the nasal and saliva samples will be done here in Kilifi. However for some test that cannot be done in Kenya, part of the samples will be sent to laboratories overseas.
We would also like to store any nasal mucus or saliva sample that is left over after we have done these tests in case they may help us to investigate new research questions. This means we are able to learn new things without taking other samples from other individuals in the future. Future research on these samples must first be approved by a national independent expert committee in Nairobi.

h) Who will have access to information about me/my child in this research?
All our research records are stored securely in locked cabinets and password protected computers. Only a few people who are closely concerned with the research will be able to view information from participants.

i) Who has allowed this research to take place?
An independent national committee and a committee in Kilifi have looked carefully at this work and agreed that the research is important, it will be conducted properly and participant's safety and rights have been respected.

j) What if I have any questions?
You may ask any of our staff questions at any time. You can also contact those who are responsible for the care of your child and this research:

PI's name(s) and contacts

Mr. Patrick Munywoki and Dr. James Nokes
KEMRI- Wellcome Trust -Kilifi District Hospital,
P.O.Box. 230, Kenya. Tel: 041 7522 063, Mobile: 0725 242233 or 0733 268290

If you want to ask someone independent anything about this research please contact
Community Liaison Manager, KEMRI – Wellcome Trust
P.O.Box 230, Kilifi. Tel: 041 7522 063, Mobile: 0723342780 or 0738472281

Or
The Secretary - KEMRI/National Ethics Review Committee
P. O. BOX 54840-00200, Nairobi, Tel: 020 272 2541 Mobile: 0722205901 or 0733400003

II. Consent form for adults

Household spread of RSV infections: Who Acquires Infection from Whom

I, _____ (name of participant)], have had the research explained to me. I have understood all that has been read and had my questions answered satisfactorily. I understand that I can change my mind at any stage and it will not affect me/child in any way.

Please insert the boxes below where relevant:
☐ I agree to take part/allow my child to take part in this research
☐ I agree to samples being stored
☐ I agree to samples being exported

Participant's signature: _____ **Date** _____
Participant's name: _____ **Time** _____
(Please print name)

I certify that I have followed all the study specific procedures in the SOP for obtaining informed consent.

Designee/investigator's signature: _____ Date _____
Designee/investigator's name: _____ **Time** _____
(Please print name)

Only necessary if the participant cannot read:

I[5] attest that the information concerning this research was accurately explained to and apparently understood by the participant and that informed consent was freely given by the participant.

Witness' signature: _____ **Date** _____
Witness' name: _____ **Time** _____
(Please print name)

Thumbprint of the participant as named above if they cannot write: _____
The participant' should now be given a signed copy to keep.

_____

[5] *A witness is a person who is independent from the study or a member of staff who was not involved in gaining the consent.*

III. Consent form for infants and under age (<18 years) children

Household spread of RSV infections: Who Acquires Infection from Whom

I, (being a parent/guardian of _____ (name of child), have had the research explained to me. I have understood all that has been read and had my questions answered satisfactorily. I understand that I can change my mind at any stage and it will not affect me/my child in any way.

Please insert the boxes below where relevant:

☐ I agree to take part/allow my child to take part in this research

☐ I agree to samples being stored

☐ I agree to samples being exported

Parent/guardian's signature: _____ **Date** _____

Parent/guardian's name: _____ **Time** _____

(Please print name)

I certify that I have followed all the study specific procedures in the SOP for obtaining informed consent.

Designee/investigator's signature: _____ **Date** _____

Designee/investigator's name: _____ **Time** _____

(Please print name)

Only necessary if the parent/guardian cannot read:

I[6] attest that the information concerning this research was accurately explained to and apparently understood by the parent/guardian and that informed consent was freely given by the parent/guardian.

Witness' signature: _____ **Date** _____

Witness' name: _____ **Time** _____

(Please print name)

Thumbprint of the parent as named above if they cannot write: _____

The parent/guardian should now be given a signed copy to keep.

---

[6] *A witness is a person who is independent from the study or a member of staff who was not involved in gaining the consent.*

Appendix H. Initial Home Visit form

I. Study on transmission of respiratory viruses within household: baseline data

Initial home visit questionnaire

Date today (DD/MM/YYYY)__ __ /__ __/ 20__ __ Time: ___:___ (in 24hrs)

Household Head details

HH head names 1_____2 _____ 3 _____

Marital status: _____

How many wives/co-wives? ____ *(Indicate number)*

Residence: Sub-location:_____

Village: _____

How many year have you spent in school? _____numbers

Household members:

How many families are in this household: _____ *(indicate number)*

How many adults (18 years and above) are in this household: _____ *(indicate number)*

How many children (under 18 years) are in this household: _____ *(indicate number)*

How many children in this household are less than five years old: _____ *(indicate number)*

How many children are in this household are less than one year old: _____ *(indicate number)*

***NB:** Please enter the details of all the household members starting with the infant in the Table overleaf*

No.	ID	Full Names	Sex (M/F)	DOB dd/mm/yyyy	Occupation	Highest level of education	Relation to the youngest infant	Live in same house as infant (Y/N)	Sleep in the same room as infant (Y/N)	Sleep in the same bed as infant (Y/N)
	10101	Alfred Katana	M	19/06/1980	Farmer	Class 7	father	Y	Y	N
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										

NB: If more than 15 continue on a separate sheet

II. Notes to the FW for completion of the Initial Home Visit form

The field workers will fill the form. It aims at collecting information about the demographic variables related to the household. Only one form is to be filled per household. It will be more appropriate if the household head (or their spouse) responds to the questions. The field worker SHOULD NOT GIVE THE RESPONDENT to fill it in by themselves.

Below is a step-by-step guide to help you administer the questionnaire effectively:

Date today and time: Note the date and time of collection of the data. The date format will be dd/mm/yyyy e.g. 28/04/2009, time, 15:30. Use the 24-hour clock system.

HH head name: Note the household head full names in the order indicated. Note this should be the name of the living household head.

Marital status: This relates to the household head and the options include married, divorced, separated, single parent and widowed. Indicate number of wives or co-wives to the HH head.

Sub-location and village: Residence details e.g. Uyombo sub-location. Possible villages include Wireless, Uyombo-centre, Uyombo-maweni, Chambuko and Madeteni.

Years spent in spend in school: Count starting from class one and indicate the number of complete years spent in school e.g. if parent dropped out of school at class 7 the years spent in school are 7 year; dropped in form 2, 10 years etc

Household members:

Household: a group people living in the same compound and with similar cooking arrangements with or without a common household head.

Family unit: Units consisting of father, mother and/or children (nuclear family unit). Indicate the number of **family units, adults** and **children** within the household

The participants' table:

Please enter the details of all the household members starting with the youngest child (herein referred as index child) in the table. This will allow easy reference when filling details on relationship and sleeping arrangement.

- Indicate the **person id** (unique study number), participant's **full names** (including **initials** of their names in brackets), **sex** and **date of birth.** For the under five children make effort to get the exact date of birth in this format day/month/ year (DD/MM/YYYY) from vaccination or clinic cards while for adults you can refer on their national identity card.
- Indicate the **occupation** of the participant e.g. Farming, fishing, housewife, house-help, teacher, casual work, business, carpenter, masonry, tapers, quarry, pupil(in primary), students(in secondary and colleges), shopkeepers, retired etc. this should be the main occupation/activity that brings income to the participant or the main activity the participant is involved in.
- Note the highest **level of education** attained or current level of education. Indicate the highest level attained by the HH head e.g. Class 7, form 3, diploma etc:
- **Relation to the** study infant: Note the relationship of all the participants in the household to the index child e.g. self, Father, Mother, brother, sister, grandfather, grandmother, stepmother, step father, uncle, aunt, househelp etc
- In reference to the index child indicate whether the participants live in the **same house**, sleep in the **same room** and /or sleep in the **same bed** as the index child.

Appendix I. Home Visit form

I. The form

Date today: ___ ___/___ ___/20_____ Time:___:___ FW Initials: _____

HH Number: ___ ___ ___ HH head name: _____ _____ _____ Visit no: _____

	StudyID	Full names	Swab taken[7] (Y/N)	Oral Fluid (Y/N)	Sick today (Y/N)	Duration Of illness (Days)	Cough (Y/N)	Runny nose (Y/N)	DIB (Y/N)	Other[8] (Y/N)	Temp (^{0}C)	Resp per min	Any treat (Y/N)	Refer (Y/N)	Comments
	10101	Alfred Katana	Y	N	Y	2	Y	Y	N	Y	36.5	28	Y	N	
1															
2															
3															
4															
5															
6															
7															

Date/time of the next appointment: ____/____/2010; Time: ____:____ hrs

---

[7] Please provide more details if nasal sample is not collected;

[8] Indicate any non-respiratory complaints at the back of this form

Additional data:

i) **Reason for not collecting the required samples** (indicate participant details, sample type and the reason)

e.g. 10101-NPS- patient not available

.....

.....

.....

.....

ii) **Other complaints** (indicate participant initials and the additional signs and symptoms)

e.g. 10101 – diarrhoea, vomiting

.....

.....

.....

.....

iii) **Treatment sought** (Indicate participant initials and the source of treatment, drugs taken especially name of any antibiotics taken)

e.g. 10101 – treatment sought from Matsangoni dispensary, amoxil, flagyl, paracetamol

.....

.....

.....

.....

iv) Any comments

.....

.....

.....

.....

.....

.....

.....

II. Notes to the FW for completion of the Home Visit form

a) Before you depart

Obtain list of households to be visited today + details of how to locate the homesteads
Fill out date of the visit at top of the home visit form
Fill in the Lab request forms for both oral and nasal swab collection
Check to take adequate materials/supplies for nasal swabbing and oral fluid collection.

b) At the household

Remember to introduce yourself to the family/mother/head of household and identify yourself as from KEMRI for the household study. Remember to always be polite and to respect the wishes of the family.

- Take care to be accurate on the Names and studyid for all the study participants

Illness assessment

- Sick today: Ask if the person has any signs and /or symptoms of respiratory illnesses on the day of the visit.
- Duration of illness: Ask when the respiratory illness started i.e. when the first signs or symptoms of respiratory illness were identified. Record duration in days e.g. Indicate 0 if only today, 1 if ill yesterday, 2 if ill day before yesterday etc. The duration should not exceed the number of days since the last home visit.
- Cough: Ask the participants (caretakers for the young children) if they have had (heard) cough today or note if you have heard the person cough during this visit.
- Runny nose/congestion: Check if the participant has running nose (common cold) or blocked nose (nasal congestion)
- DIB (Difficulty in breathing): Check participants for difficulty in breathing.
- Other complaints: especially non-respiratory symptoms: Indicate any other complaints at the back of the form indicating the duration of the symptoms.
- Temp: Take axillary temperature and record here.
  Respiratory rate: Also referred as breaths per minute. Record the breathing rate as accurately as possible. Ensure the participant has settled before recording. Assess if the participant has fast breathing (i.e. >50 for all ages, and >40 if 12 months or over). This will be collected only for under 5 years old or persons with respiratory illness.
- Any treatment: Did you seek any treatment for this illness? Provide details of the treatment sought in the box provided. The details should include the type of drugs taken, source of the treatment, when the treatment was sought.
- Refer: Record here if the participant was referred to either Matsangoni dispensary. If child has fast breathing for age OR difficulty in breathing AND cough OR nasal discharge/congestion), then refer the child to the Matsangoni dispensary for further review and treatment (complete referral note in RSV exercise book). Also refer if the temperature >=37.5°C to get a malaria blood slide. Emphasise that the mother should take the child today.
- Any comments: Any other findings/observations. Indicate more details in the box provided

Specimen collection

- NPSwab: Record if nasopharyngeal flocked swab has been collected. If not taken indicate the reason.
- Oral fluid: Record if saliva sample was taken. If not taken indicate the reason.

c) Note: Plan for the next appointment (within 3 to 4 days)

Appendix J. Clinic Visit Form

I. The form

Date today (DD/MM/YYYY) __ __ /__ __/ 20__ __ Time: ___:___hrs

a) Person details

Patient Names: 1_____ 2_____ 3_____

Study Person id: __ __ __ __ __ __ __

DSS PID: __ __ __ __ __ __ __ __

Sex: ____ (M/F)

Residence (Village): _____

Brought to clinic by (caretaker names): _____

Relation to the patient: _____

Referred by: _____FW/Self/Study clinician

b) Illness assessment

Anthropometrics:

Weight	[ ][ ] [ ].[ ][ ] kg	To be filled by the FW
MUAC	[ ] [ ].[ ] cm	
Height/Length	[ ] [ ] [ ].[ ] cm	
Oxygen saturation	[ ] [ ] [ ] %	
Heart rate	[ ] [ ] [ ]/min	
Axillary temperature	[ ] [ ].[ ] ^{0}C	
Respiratory rate	[ ] [ ]/min	
Is the participant sick today?	[ ] y/n	
How long has the patient been sick?	[ ] [ ] days	

What are the main complaints? (Please indicate Yes or No)

Fever	[ ] y/n	Crackles	[ ] y/n
Cough	[ ] y/n	Wheezes	[ ] y/n
Nasal discharge/congestion	[ ] y/n	Unable to feed	[ ] y/n
Difficulty in breathing	[ ] y/n	Head nodding	[ ] y/n
Nasal flaring	[ ] y/n	Cyanosis	[ ] y/n
Sore throat	[ ] y/n	Unable to talk in complete sentences	
Chest wall indrawing	[ ] y/n	[ ] y/n	

Conscious level (please tick one)

Alert ☐ Lethargic ☐

Prostrate ☐ Unconscious ☐

i) Other complaints (specify in the box below)

ii) Laboratory tests:
iii) Malaria test: Indicate results as either positive, negative, equivocal or not done
iv) Rapid diagnostic test: [ ]
v) Peripheral blood slide: [ ]
vi) HIV test done: [ ] y/n
vii) Primary diagnosis (specify): _____
viii) Other diagnoses (specify in the box below)

ix) Treatment given (List all the drugs prescribed in the box below)

x) Does the child require referral for hospital (KDH) admission [ ] (y/n)
xi) Initials of the clinician reviewing the patient [ ]

c) Specimen Collection

i) Does the patient require a nasal sample? [ ] (y/n)
ii) If yes, has the swab been collected: [ ] (y/n)
iii) If nasal sample is not collected specify reason: _____
iv) Nasal specimens collected by (indicate your initials) [ ]

II. Notes for the completion of the Clinic visit Form

a) General instructions

The form will be filled by the clinician but could be assisted by the study field worker. This clinical assessment form is to be filled for all clinic visits made by the household study participants presenting at Matsangoni Dispensary or KDH outpatient with any illness Confirm the participant's particulars from their study cards/book.

b) Patient details

Indicate the patient full names, the person id (unique study number) and the DSS PID as indicated in the study card/book.

- Sex: Indicate whether male or female
- Residence: This should be the village of residence

- Caretaker: Refers to the person who brought or accompanied the patient to the clinic. Note the relation of the caretaker with the patient. He/she could be from the same household as the patient or not.
- Referral: Establish who referred the patient to the clinic. Whether it's the study FWs, clinician or it's a self referral

c) Illness assessment

- Weight: measured in Kgs into 2 decimal places. Ensure the weighing scales are calibrated every week
- MUAC: mid-upper arm circumference; reported in cm. This will only be collected in children who are below 13 years old.
- Height/Length: measure in cms
- Oxygen saturation: Use pulse oximeter to measure oxygen saturation. Recheck if the oxygen saturation is below 95% and if it is less than 50% there is likely to be a problem with the Oximeter – repeat measurement. If less than 90% repeat the procedure and compare the two readings. Ensure person is at rest and at ease.
- Heart rate: or pulse rate: Use the readings from the pulse oximeter or count peripheral pulse rate for one complete minute.
- Axillary temperature: Take axillary temperature using a digital thermometer e.g. accuracy 37.3
- Respiratory rate: Also referred to as breaths per minute. Count respirations for one complete minute. Ensure person is in calm stated and rested before starting. If a child it should not be agitated or crying for example. Record the breathing rate as accurately as possible. Ensure the participant has settled before recording. Assess if child has fast breathing (i.e. >50 for all ages, and >40 if 12 months or over).
- Sick today: Ask if the participant has any illness today.
- Duration sick: note the duration the patient has been ill in days. Indicate 0 if only today, 1 if ill yesterday, 2 if ill day before yesterday etc
- Main complaints: during the clinical assessment of the illness establish whether any of the listed respiratory-related signs and symptoms are reported.
- Other complaints: Note the presence of any other complaints which might be accompany respiratory illness e.g. diarrhoea, vomiting
- Laboratory tests: Please indicate any malaria parasite test results for Rapid diagnosis and or blood slide as positive, negative, equivocal or not done. Establish if a HIV test has been done.
- Primary diagnosis: Indicate the final/main diagnosis based on patient complains and any laboratory tests done
- Other diagnoses: Note any other diagnosis where applicable.
- Referral: Indicate if the child requires further clinical care (hospital admission) in the Kilifi District Hospital.

d) Specimen Collection

Assess if the patient requires a nasal specimen. If the last sample was collected more than 2 days ago please take a nasal swab. If nasal sample is not collected specify reasons. Please indicate the initials of the FW/clinician collecting the sample.

Appendix K. The Risk Survey Questionnaire

I. Baseline data

Date today (DD/MM/YYYY)__ __ /__ __/ 20___

Household number: [ ]

Household name: _____

II. Demographic, social-economic and environmental factors

i) Where does the index child normally remain during the day? *Please tick one*

Inside the home	
Outside home	

ii) Who usually takes care of the index child during the day? *Please tick one*

Mother	
Another family member	
House help	
Other[ ]	

iii) Which of the following do you have in the household at the present time? *Please tick as appropriate*

Bicycle		Radio	
TV/video deck		Mobile phone	
Motorcycle		Electricity	

iv) How many animals do you have?

Cows		Sheep		Turkey	
Goats		Chicken		Others [ ]	

v) What type of main house does the principal care taker live in? *Please tick one*

Mud wall house (Makuti thatching)	
Mud wall house (Iron sheets roof)	
Block wall house(Makuti thatching)	
Block wall house(iron sheets)	

vi) Is this house.....? *Please tick one*

Owner occupied		No rent with consent of owner	
Rented		Other [ ]	

vii) What type of toilet do you have? *Please tick one*

Flush toilet inside		Latrine inside		No toilet	
Flush toilet outside		Latrine outside			

Into dwelling	
In compound or plot	
public	

viii) Do you burn your refuse (waste) within the homestead?

Yes	
No	

ix) What type of fuel do you use for cooking? *Please tick as appropriate*

Fire wood		Gas		Other [ ]	
Wood Charcoal		Paraffin			

x) Where do you do your cooking? *Please tick as appropriate*

Same house as you sleep in	
Different house from where you sleep	
Outside	

xi) Where do you get water for domestic use? *Please tick as appropriate*

Piped		Open well	
Surface e.g rivers, swamps		Closed well	

xii) Site of piped or well water: *Please tick one*

III. Summary table

No.	*StudyID*	*Full names*	*Date of Birth*	*In school*		*Smokes (y/n)*	*Weight(kg)*	*Height (m)*	*MUAC (Left arm)*	*Remarks*
				(y/n)	*Class*					

Appendix L. Nasopharyngeal flocked swab procedure

The standard operating procedure

I. Introduction

This SOP describes the nasal flocked swab procedure for collecting nasopharyngeal epithelial cells for viral and bacterial analysis. This procedure has previously been considered more acceptable than other nasal specimen collection methods.
The Nasal flocked swab is simple procedure, which entails use of commercially available flocked swabs in collection of nasal mucous specimen. A flocked swab is a plastic shaft with a tip covered with nylon fibres that acts like a soft brush. It is the responsibility of the RSV household field workers to do this procedure regularly under training and supportive supervision from the project coordinator and the PI.

II. Purpose

This SOP describes the nasal flocked swab procedure for collecting nasopharyngeal epithelial cells for viral and bacterial analysis.
This procedure has previously been considered more acceptable than other nasal specimen collection methods.

III. Scope

This SOP targets designated RSV household Field Workers.

IV. Responsibility

It is the responsibility of the RSV household field workers to do this procedure regularly under training and supportive supervision from the project coordinator and the PI.

V. Definitions

The Nasal flocked swab is simple procedure, which entails use of commercially available flocked swabs in collection of nasal mucous specimen. A flocked swab is a plastic shaft with a tip covered with nylon fibres that acts like a soft brush.

VI. Abbreviations and Terms:

- RSV – respiratory syncytial Virus
- NPS– Nasopharyngeal Flocked Swab
- FW– Field Worker
- SOP – Standing Operating Procedure
- QA – Quality Assurance
- PI – Principal Investigator

VII. Specimen

Nasopharyngeal epithelial cells.

VIII. Equipment/ materials/ reagents

- RSV Household study laboratory request form (attached)
- Gloves
- Pliers
- Sterile nasopharyngeal flocked swabs
- Specimen tubes with 1ml of Universal transport medium
- Indelible marker pen
- Cool box with ice packs

IX. Procedures:

a) Preparation:

- Ensure all materials required are available
- Confirm that the participant (>18 years) and or parent/guardian (if child) has signed informed consent form for the RSV household study.
- Explain the procedure to the participant or parent/guardian (if child) demonstrating the equipment and the materials to be used.
- Reassure the participant in case they are anxious that the procedure will cause very mild discomfort and will be quick (~10 secs)

b) Methodology

1. Wash hands and wear clean gloves
2. Label specimen bottle with study Identification number, name, and date.
3. Hold person's head firm and steady (if infant or child-immobilize arm)
4. Measure the distance between the patient's nares and earlobe to estimate the length of insertion.
5. Insert the flocked swab into the nostril towards the pharynx until a slight resistance is met.
6. Rotate the swab three times to obtain epithelial cells and hold the swab in place for 5 seconds to ensure maximum absorbency.
7. Withdraw the swab gently and put in 2ml of transport medium and break the shaft at the painted breakpoint.
8. Place specimen container in the cool box immediately
9. Dispose off the gloves and wash hands
10. Fill in the sample form – remember to include time of collection
11. Ensure the sample is transported to the Lab as soon as possible where it will be stored at 4°c.

NB: Do not delay in getting sample to a 4^0C fridge. If cannot get to Fridge immediately, store in cool box temporarily.

Appendix M. Oral fluid sample collection procedure

The standard operating procedure

I. Introduction

This SOP describes the oral fluid sampling procedure for collecting gingivo-crevicular fluid from between the gums and teeth. This procedure has previously been considered more acceptable than other specimen collection methods particularly in non-clinical settings. This SOP targets designated field workers for the study on RSV transmission within households. It is the responsibility of the field workers to do this procedure regularly under training and supportive supervision from the project coordinator and the PI.
Oral fluid sampling is a simple procedure, which entails use of commercially available devices to collect gingivo-crevicular fluid from between the gums and teeth.

II. Purpose

This SOP describes the oral fluid sampling procedure for collecting gingivo-crevicular fluid from between the gums and teeth
This procedure has previously been considered more acceptable than other specimen collection methods.

III. Scope

This SOP targets designated field workers for the study on RSV transmission within households.

IV. Responsibility

It is the responsibility of the field workers to do this procedure regularly under training and supportive supervision from the project coordinator and the PI.

V. Definitions

Oral fluid sampling is a simple procedure, which entails use of commercially available devices to collect gingivo-crevicular fluid from between the gums and teeth.

VI. Abbreviations and terms

- RSV – respiratory syncytial Virus
- OF – Oral fluid
- FW– Field Worker
- SOP – Standing Operating Procedure
- QA – Quality Assurance
- PI – Principal Investigator

VII. Specimen

Gingivo-crevicular fluid

VIII. Equipment/ materials/ reagents

- Lab request form (attached)
- Disposable gloves and Paper towel
- ORACOL collection device (Malvern Medical Developments, Worcester, UK)
- Labels
- Indelible marker pen
- Cool box with ice packs

IX. Procedures:

a) Preparation:

- Ensure all materials required are available
- Confirm that the participant (>18 yrs) and or parent/guardian (if child) has signed informed consent form for the RSV household study.
- Explain the procedure to the participant or parent/guardian (if child) demonstrating the equipment and the materials to be used.
- Reassure the participant in case they are anxious that the procedure will not cause any discomfort and will last for about a minute.

b) Methodology

1. Wash hands and wear clean gloves
2. Label specimen tube with study Identification number, name, and date.
3. Remove ORACOL device from covering and transport tube
4. Holding by stick rub sponge gently but firmly around gums (front and back, upper and lower, including base) for 1 minute.

IMPORTANT: The sponge must be fully saturated with oral fluid and so one minute of use is required.

5. Replace device into stoppered tube (sponge at bottom).
6. Dispose of gloves and wash hands.
7. Place specimen container in the cool box immediately
8. Dispose off the gloves and wash hands
9. Fill in the sample form – remember to include time of collection
10. Ensure that the sample is transported to the Lab as soon as possible, where it will be stored at 4°c.

NB: Do not delay in getting sample to a 4^0C fridge. If cannot get to Fridge immediately, store in cool box temporarily.

X. Attachment: Household study specimen collection form

Household study: LAB REQUEST FORM

Partcipant's name: _____ DATE __/__/__

Study No. ☐☐☐☐☐ TIME __ __

Other details _____

Sample details (please tick)

Flocked Swab ☐ Oral Fluid ☐

Collected by (initials) ☐

FOR LAB USE ONLY

Results(Please indicate appropriately ; Postive, negative or equivocal)

Done By(initials) ☐

Appendix N. Real time multiplex PCR (M-PCR) detection of respiratory viruses

The Standard Operating Procedure

I. Introduction

Influenza A and B are important viral infections of the respiratory tract of adults and children. Most cases occur during the annual winter epidemics. Severe infections can occur in the immunocompromised, those with chronic cardiac, pulmonary or metabolic disease and in the extremes of age. Most symptoms include fever, myalgia, sore throat and cough. In children, gastrointestinal symptoms also are present. Complications of influenza infection include primary (viral) or secondary (bacterial) pneumonia, cardiac involvement and neurological illness (including encephalopathy, encephalitis, Reyes syndrome and Guillain-Barre Syndrome.

Influenza C is a common cause of mild or asymptomatic respiratory infection. However, please note that testing has not been common in the past and therefore we may find this pathogen in a wide range of previously unrecognised clinical situations.

Parainfluenza viruses are common causes of upper and lower respiratory tract infection. Types 1 and 2 tend to occur mostly in the winter months whereas type 3 occurs mainly in the spring. It is unclear if parainfluenza 4 has a defined seasonality. Type 1 and 2 are commonly associated with croup in young children. Type 3 is second only to RSV as a cause of bronchiolotis and pneumonia in infants. Type 4, like influenza C, is thought to be a common cause of mild or asymptomatic respiratory infection. Please note that these pathogens can also cause severe infection in immunocompromised patients, especially type 3.

Rhinovirusescause ~one third of colds. They are common in all age groups and can cause upper and lower respiratory tract infection. Increased testing has implicated these viruses in severe infections especially in those with asthma and COPD. Severe infections are also likely to occur in the immunocompromised. There are over 100 types of rhinovirus known to cause respiratory infection in humans.

Coronaviruses 229E, OC43 and NL63 are frequent causes of the common cold (2^{nd} only to rhinoviruses). There have also been reports of pneumonia in the elderly, infants and immunocompromised. Outbreaks of pneumonia have also been described in healthy adults. Increased testing will probably lead to detection of coronaviruses in severe upper and lower respiratory tract infection in other patient groups.

Respiratory syncytial virus is a common cause of respiratory illness in all age groups. Most infections tend to occur during the winter season (December-March) in temperate areas, and with the rainy seasons in the tropics. RSV is of particular importance as a cause of severe lower respiratory tract infection in infants (causing bronchiolitis and pneumonia), the immunocompromised (especially BMT patients) and the elderly.

Adenovirus is a common cause of respiratory infection (ranging from pharyngitis to severe pneumonia).

Mycoplasma is a bacterial pathogen recognised to be a cause of atypical pneumonia in all age groups. PCR has shown to be a sensitive alternative to culture and serology.

Five triplex RT-PCR reactions are performed to detect the following pathogens:

- Set 1 - Influenza A, B and C viruses
- Set 2 - Parainfluenza viruses 2, 3 and 4
- Set 3 - Adenovirus, Respiratory syncytial viruses A and B
- Set 4 - Coronaviruses 229E, OC43 and NL63
- Set 5- Parainfluenza virus 1 and human metapneumovirus (hMPV) + Mycoplasma Pneumonia
- Set 6- Single RT-PCR Reaction to detect Rhinovirus

II. Scope / Responsibility

Viral Epidemiology and Control (VEC) lab personnel

III. Definitions

- PC - Positive control.
- NTC - Non template control.
- NC - Negative control.
- NA - Nucleic acids.
- VEC- Virology epidemiology cluster.
- RNA - Ribonucleic acid.
- DNA - Deoxyribonucleic acid.
- CDNA - Complimentary deoxyribonucleic acid.
- NPA- Nasal pharyngeal aspirate
- GW - Gastric washing.

IV. Safety/ Risk Assessment:

1. Avoid any loose disconnection of electrical wire that can cause shock.
2. All respiratory samples are potentially infectious; ensure all proper protective equipment are used.
3. All work should be done in safety cabinets

V. Specimen:

Samples submitted are likely to be RNA/DNA from respiratory samples (throat/nasal swabs, NPA, sputum). In some cases other sample types may be tested including post mortem material (heart and lung), induced sputum, oral fluid, GW, and CSF.

VI. Materials:

a) Equipment

- ABI Prism 7500
- ABI Prism reaction plates
- ABI Prism reaction plate seals
- Pipettes for 1-1000µl
- Tips for 1-1000µl
- Multichannel Pipettes

b) Reagents: RT-PCR Reactions

- Quantifast Multiplex RT-PCR kit
- Respiratory pathogen triplex-set mixes oligonucleotides (attached)
- FLU (Influenza A, B and C) primer and probe mix
- PIV (Parainfluenza virus types 2, 3 and 4) primer and probe mix
- RSV (Adenovirus, RSV-A, RSV-B) primer and probe mix
- HCoV (Coronavirus 229E, OC43 and NL63) primer and probe mix
- OTHER (Parainfluenza Type 1, Human Metapneumovirus A + B and Mycoplasma Pneumonia) primer and probe mix.
- RHINO (Rhinovirus primer and probe mix)
- Nuclease free or Molecular Biology grade water
- The positive control material can be either pooled cultured virus, extracted RNA or PCR product. Working stocks of each control, diluted so they will have a Ct between 25 and 30 in a standard reaction, are stored in single target aliquots at -20 °C. One control to be included for each viral target.

VII. METHODOLOGY:

a) Principle

Viral RNA is transcribed into cDNA using a specific primer mediated reverse transcription step followed immediately in the same tube by polymerase chain reaction. Detection of product is via a dual labeled molecular probe for each virus of the triplex PCR.

b) PCR procedure

1. Prepare the RT-PCR Master Mix and Primer/Probe Mixes correctly (see table below) for preparation procedure.
2. For stored NA remove the sample plates from the -80°C freezer and leave to defrost on ice in the PCR template addition room. For fresh NA place on ice and proceed to the master mix room.
3. Remove all the tubes from a Quantifast Multiplex RT-PCR kit and the aliquots of primers and probe needed for the PCR from -20 °C freezer and defrost on ice. Move everything to the PCR master mix room.
4. Make up the master mix for the PCR according to table 1. Always make up enough master mix to allow for controls and pipetting errors. For example, for a full 96-well plate make up enough master mix for 110 reactions.
5. Vortex the master mix and then pipette 8ul of the PCR mix into the 96-well plate.
6. Move to the template addition room and add 2ul of the NA from the Sample Plate to the PCR plate. At each stage visually check the volume of liquid in the pipettes before adding it to the PCR plate.

Table 1: Mastermix volumes for the triplex reactions.

Reaction	Reagent	1rxn(µl)	12rxn(µl)	50rxn(µl)	110rxn(µl)
RSV Corona Flu PIV	2x Quantifast	5.0	60	250	550
	50x ROX	0.2	2.4	10	22
	Primer/Probe Master	1.5	18	75	165
	Quantifast RT	0.1	1.2	5.0	11
	Water	1.2	14.4	60	132
Other	2x Quantifast	5.0	60	250	550
	50x ROX	0.2	2.4	10	22
	Primer/Probe Master	2.0	24	100	220
	Quantifast RT	0.1	1.2	5.0	11
	Water	0.7	8.4	35	77
Rhino	2x Quantifast	5.0	60	250	550
	50x ROX	0.2	2.4	10	22
	Primer/Probe Master	0.5	6.0	25	55
	Quantifast RT	0.1	1.2	5.0	11
	Water	2.2	26.4	110	242

7. In the general lab bench add 2ul of PC and NTC in respective wells. For each plate there should be an individual control for each target and then a control where all three targets are mixed in one well.
8. Seal the PCR plate with an optical sealing sheet, ensuring all the edges are firmly sealed to prevent possible evaporation in the thermocycler. Pulse vortex, spin and load the plate to the machine.
9. On the ABI 7500, open up the RV_MPX template and either add your sample ID's manually or import from the library save and start run.
10. Once the PCR has finished analyze, record and report the results to the appropriate people. (for analysis see attachment 3)

VIII. Attachments:

- List of the primer and probe sequences
- Guide to preparing the RT-PCR Master Mix aliquots and Triplex-mix aliquots
- Results analysis
- Sample Extraction Sheet
- Sample Plate Sheet

a) Attachment 1: Primer and probe sequences

Virus Target	Target Gene	Forward primer sequence	Reverse primer sequence	Probe sequence	Product size (bp)
Influenza A	Matrix	AAGACAAGACCAATYCTG TCACCTCT	TCTACGYTGCAGTCCYCGC T	**FAM**-TYACGCTCACCGTGCCCA GTG-**BHQ1**	93
Influenza B	NS	ATGATCTTACAGTGGAGG ATGAAGAA	CGAATTGGCTTTGRATGTC CTT	**CY5**-ATGGCCATCGGATCCTCAA YTCACTCT-**BHQ2**	91
Influenza C	Matrix	GGCAAGCGACATGCTGAA YA	TCCAGCTGCYTTCATTTGC TTT	**VIC**-CTCTTCCTTCTGATTTTTTC AAA-**MGBNFQ**	85
Human metapneumo virus A	Fusion	GCYGTYAGCTTCAGTCAAT TCAA	TCCAGCATTGTCTGAAAAT TGC	**VIC**-CAACATTTAGAAACCTTCT-**MGBNFQ**	69
Human metapneumo virus B	Fusion	GCYGTYAGCTTCAGTCAAT TCAA (Common with A)	GTTATCCCTGCATTGTCTG AAAACT	**VIC**-CGCACAACATTTAGGAATC TTCT-**MGBNFQ**	74
Parainfluenza virus 1	HN	GTGATTTAAACCCGGTAAT TTCTCA	CCTTGTTCCTGCAGCTATT ACAGA	**FAM**-ACCTATGACATCAACGAC-**MGBNFQ**	83
Parainfluenza virus 2	HN	ATGAAAACCATTTACCTAA GTGATGGA	CCTCCYGGTATRGCAGTGA CTGAAC	**VIC**-TCAATCGCAAAAGC-**MGBNFQ**	68
Parainfluenza virus 3	HN	CCAGGGATATAYTAYAAA GGCAAAA	CCGGGRCACCCAGTTGTG	**FAM**-TGGRTGTTCAAGACCTCCA TAYCCGAGAAA-**BHQ1**	101
Parainfluenza virus 4	Fusion	CAGAYAACATCAATCGCC TTACAAA	TGTACCTATGACTGCCCCA AARA	**CY5**-CCMATCACAAGCTCAGAA ATYCAAAGTCGT-**BHQ2**	80
Human coronavirus 229E	Nucleocapsid	CAGTCAAATGGGCTGATG CA	AAAGGGCTATAAAGAGAA TAAGGTATTCT	**FAM**-CCCTGACGACCACGTTGT GGTTCA-BHQ1	76
Human coronavirus OC43	Nucleocapsid	CGATGAGGCTATTCCGACT AGGT	CCTTCCTGAGCCTTCAATA TAGTAACC	**CY5**-TCCGCCTGGCACGGTACTC CCT- **BHQ3**	76
Human coronavirus	1a gene	ACGTACTTCTATTATGAAG CATGATATTAA	AGCAGATCTAATGTTATAC TTAAAACTACG	**VIC**-ATTGCCAAGGCTCCTAAAC GTACAGGTGTT-**TAMRA**	103

NL63					
Respiratory syncitial virus A	NP	AGATCAACTTCTGTCATCC AGCAA	TTCTGCACATCATAATTAG GAG	**FAM**-CACCATCCAACGGAGCAC AGGAGAT-**BHQ1**	84
Respiratory syncitial virus B	NP	AAGATGCAAATCATAAAT TCACAGGA	TGATATCCAGCATCTTTAA GTA	**VIC**-TTTCCCTTCCTAACCTGGAC ATA-**TAMRA**	103
Rhinovirus	5'-UTR	TGGACAGGGTGTGAAGAG C	CAAAGTAGTCGGTCCCATC C	**VIC**-TCCTCCGGCCCCTGAATG-**T AMRA**	144
Adenovirus	Matrix	GCCACGGTGGGGTTTCTAA ACTT	GCCCCAGTGGTCTTACATG CACATC	**CY5**-TGCACCAGACCCGGGCTCA GGTACTCCGA-**BHQ2**	132
Mycoplasma pneumoniae	Cytadhesin P1 (P1) gene	AAGCAGGAGTGACGGAAA CAC	CACCACATCATTCCCCGTA TT	**CY5**-CTCCACCAACAACCTCGCG CCTA-**BHQ2**	73

b) Attachment 2: Preparation of RT-PCR mastermix and primer-probe mixes

The primers and probes will arrive from the manufacturers either lyophilised or at a concentration of 100uM. Lyophilised primers or probes will come with details on how to re-suspend them to a concentration of 100uM. This should be done with molecular grade PCR water or TE buffer. All re-suspended oligos should be stored at -20 until needed. Freeze thaw cycles should be kept to an absolute minimum to avoid degradation of the primers and probes. A working concentration of 50uM should be created from the 100uM stock. This can achieved by adding 250ul of Master stock (100uM) to 250ul of molecular grade PCR water or TE buffer. The purpose of the working stock is to avoid repeated freeze-thaws of the master stock.

The working concentration of all primers is 5uM and all probes is 2.5uM. The oligos should be diluted in molecular grade PCR water or TE buffer from the working stock (20uM) using the following protocol:

RSV, PIV, FLU, Corona (final volume of 500ul)

5mM	Concentration	Volume	Final concentration
Water	n/a	125ul	n/a
RSV A* fwd	50uM	50ul	5mM
RSV A rev	50uM	50ul	5mM
RSV A prb	50uM	25ul	2.5mM
RSV B* fwd	50uM	50ul	5mM
RSV B rev	50uM	50ul	5mM
RSB B prb	50uM	25ul	2.5mM
Adeno* fwd	50uM	50ul	5mM
Adeno rev	50uM	50ul	5mM
Adeno prb	50uM	25ul	2.5mM

Other (final volume of 500ul)

Water	n/a	0ul (no water needed)	n/a
PIV 1 fwd	50uM	50ul	5mM
PIV 1 rev	50uM	50ul	5mM
PIV 1 probe	50uM	25ul	2.5mM
hMPV A fwd	50uM	50ul	5mM
hMPV A rev	50uM	50ul	5mM
hMPV A prb	50uM	25ul	2.5mM
hMPV B fwd	50uM	50ul	5mM
hMPV B rev	50uM	50ul	5mM
hMPV B prb	50uM	25ul	2.5mM
M.p fwd	50uM	50ul	5mM
M.p rev	50uM	50ul	5mM
M.p prb	50uM	25ul	2.5mM

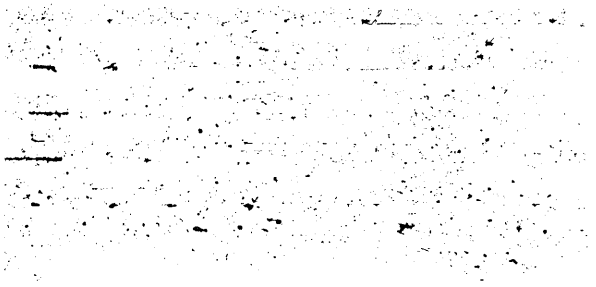
Rhino (final volume of 500ul)

Water	n/a	425ul	n/a
Rhino fwd	50uM	50ul	5mM
Rhino rev	50uM	50ul	5mM
Rhino prb	50uM	25ul	2.5mM

*In this set you should replace these listed primers for those relevant to the triplex reaction involved.

c) Attachment 3: Results Analysis

As part of quality assurance results should be double-checked.
This is done by dragging the "Baseline End" icon manually to the required position.
You can also adjust the threshold by dragging if you wish to remove some 'late noise' positive that have crept above the threshold. However, if you are unsure about any possible positive then it is a good idea to view the component graph for that well. This may give you a more accurate picture of whether there was amplification in that well or not. Compare the component traces of your unknown samples against the control well or any other positive sample.
The graph axis can also be altered to get a more accurate placement of your threshold level. Some reactions will result in higher fluorescence than other leading to difficultly analysing the lower level reactions. Manually alter the y-axis scale in the graph settings tabs to solve this problem.
The results should then be saved in monthly folders.
Run manual back up.
Export your results to a folder and send by mailing.
Upload results to file maker.
This sheet should **ALWAYS** be doublechecked.



d) Attachment 4: Sample extraction sheet

Sample Extraction Sheet#		Operator		Date		Database Completed	
Extraction Kit Batch#		-80 Freezer Box#					

EXTRACTIONS:

Tube Number	Sample ID		Tube Number	Sample ID		Tube Number	Sample ID	
1			9			17		
2			10			18		
3			11			19		
4			12			20		
5			13			21		
6			14			22		

e) Attachment 5: Sample plate sheet

RESPIRATORY VIRUS MULTIPLEX MASTER SAMPLE PLATES SHEET

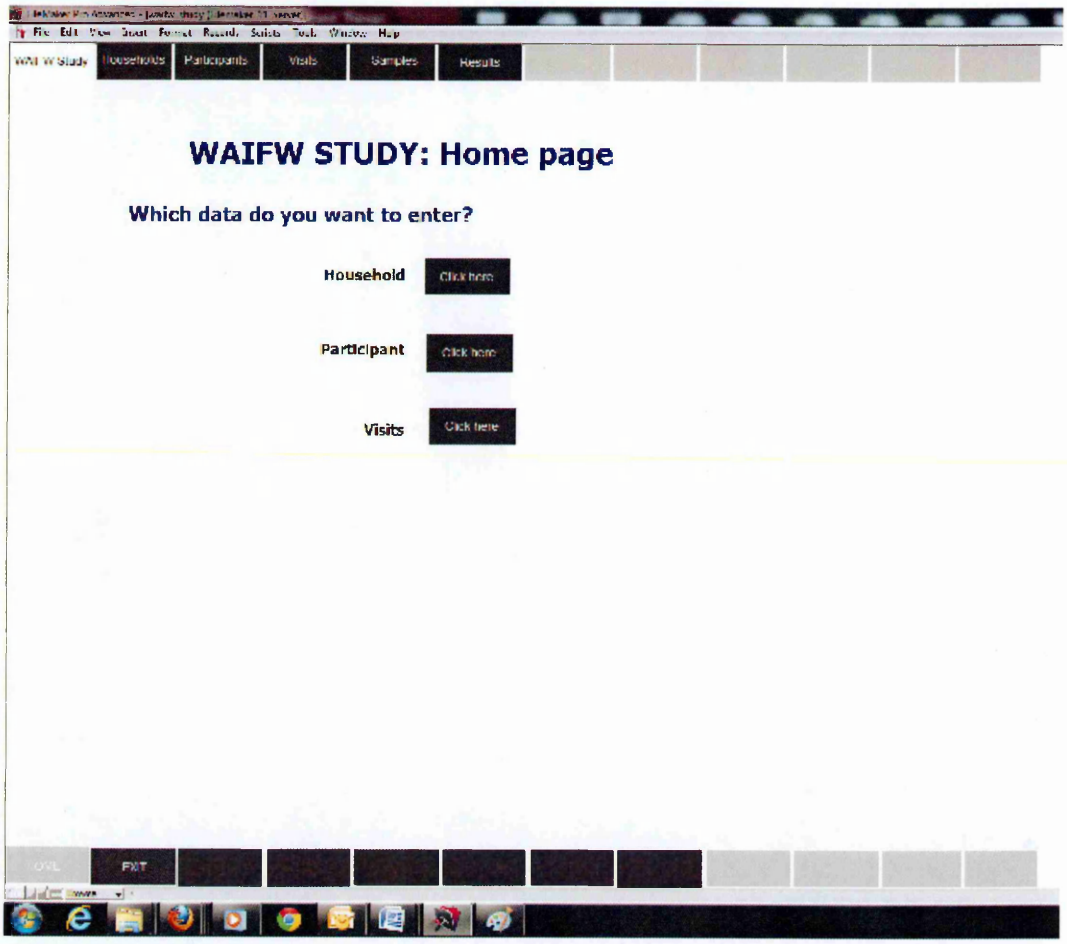
Sample Plate Number		Operator		Date		Database completed	

SAMPLE ID:

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12

Appendix O. Database screen shots

a) The database home page

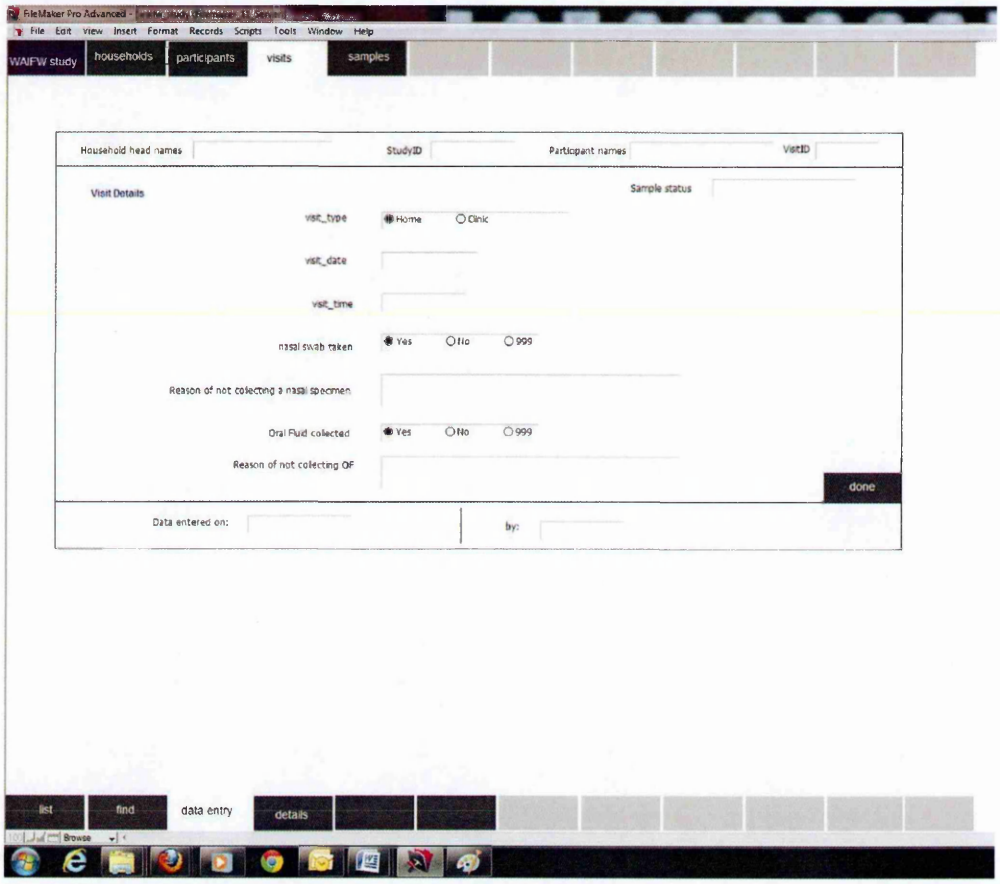


A screenshot of the database homepage. Layouts for entering sample details (freezer storage positions and laboratory results) were accessed by clicking the 'Visits' tab in the homepage. There were several alternatives to this when in other layouts.

b) Table relationships

The database has four main tables with one-to-many relationship. The tables contain household, participant, visits and samples variables (fields)

c) Sample layout for data entry



Appendix P. Scientific and ethical approval letters

I. Approval letter from Biomedical Research Ethics Sub-Committee, Warwick Medical School, UK

8th September 2009

centre for primary health care studies

Dr James Nokes
c/o KEMRI/WT Research Programme
PO Box 230
Kilifi 80108
Kenya
East Africa

Dear James

Project Title:
Transmission of RSV within the household

Thank you for submitting your revisions for the above-named project to the University of Warwick Biomedical Research Ethics Sub-Committee for Chair's Approval.

I am pleased to confirm that the revised documentation meets the required standard. I am also able to confirm that BREC has received confirmation from the University's insurance office that cover is in place under the University's Professional Indemnity insurance and that the study will be sponsored by the University of Warwick. This means that the above study has been granted full approval by BREC allowing the study to commence as confirmed in our 8th September 2009 email to you.

May I remind you any substantial amendments require approval from the Committee and that, once your study is completed, the Committee should receive an End of Project Report.

Yours sincerely,

Professor Jane Barlow
Chair
Biomedical Research
Ethics Sub-Committee

Copy:
Lynn Green, Research Governance Facilitator, WMS, University of Warwick
Mrs Carol Howes. Executive Officer, Biological Sciences, University of Warwick
David Bennett, Senior Research Manager, WMS, , University of Warwick

Warwick
Medical School

Warwick Medical School
Coventry CV4 7AL United Kingdom
General Enquiries: +44 (0)24 7657 2950
Fax: +44 (0)24 7652 8375
Website: http://www.warwick.ac.uk/primary_care

II. KEMRI-SSC approval letter

KENYA MEDICAL RESEARCH INSTITUTE

P. O. Box 54840 - 00200 NAIROBI, Kenya
Tel: (254) (020) 2722541 , 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: kemri-hq@nairobi.mimcom.net; director @ kemri. org; website. www.kemri.org

ESACIPAC/SSC/4766 — 14th July, 2009

DIRECTOR
CENTRE FOR ... RESEARCH, CGMR-C

KEMRI / WELLCOME TRUST
RESEARCH PROGRAMME
20 JUL 2009
RECEIVED
P. O. Box 230-80108 KILIFI

Patrick Munywoki

Thro'
Director, CGMR-C
P.O. Box 428
KILIFI

forwarded 20/7/09

REF: **SSC No. 1651 (Revised) – Household transmission of respiratory syncytial virus (RVS): who acquires infection from whom**

I am pleased to inform you that the above-mentioned proposal, in which you are the PI, was discussed by the KEMRI Scientific Steering Committee (SSC), during its 158th meeting held on 30th June, 2009 and has since been approved for implementation by the SSC.

The SSC however, advises that work on this project can only start when ERC approval is received.

for C. Mwandawiro, PhD
SSC SECRETARY

III. Ethical review letters

a) First approval letter

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840 - 00200 NAIROBI, Kenya
Tel: (254) (020) 2722541 , 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: kemri-hq@nairobi.mimcom.net; director @ kemri. org; website: www.kemri.org

KEMRI/RES/7/3/1 **September 16, 2009**

TO: **MR. PATRICK MUNYWOKI (PRINCIPAL INVESTIGATOR)**

THRO': **DR. NORBERT PESHU,**
THE DIRECTOR, CGMR-C,
KILIFI

Forwarded 21/9/09

KEMRI/ WELLCOME
RESEARCH PR
22 SEP 2009
RECEIVED
P.O. Box

RE: **SSC No. 1651(*ETHICS REVIEW*): HOUSEHOLD TRANSMISSION OF RESPIRATORY SYNCYTIAL VIRUS (RSV): WHO ACQUIRES INFECTION FROM WHOM?**

Make reference to your letter dated September 1, 2009. This is to inform you that the issues raised during 169th meeting of KEMRI/National Ethics Review Committee held on Tuesday 18th August 2009, have been adequately addressed.

Due consideration has been given to ethical issues and the study is hereby granted approval for implementation effective this **16th day of September 2009,** for a period of twelve (12) months.

Please note that authorization to conduct this study will automatically expire on **Wednesday, 15th September 2010.** If you plan to continue with data collection or analysis beyond this date, please submit an application for continuing approval to the ERC Secretariat by **Wednesday, 4th August 2010.**

You are required to submit any amendments to this protocol and other information pertinent to human participation in this study to the ERC prior to initiation. You may embark on the study.

Yours sincerely,

R.C. Kithinji

R. C. KITHINJI,
FOR: SECRETARY,
KEMRI/NATIONAL ETHICS REVIEW COMMITTEE

b) ERC annual ethical approval

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840 - 00200 NAIROBI, Kenya
Tel: (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

DIRECTOR CGMR-C KEMRI

KEMRI/RES/7/3/1 **September 10, 2010**

TO: **DR. PATRICK MUNYWOKI,**
PRINCIPAL INVESTIGATOR

THRO': **DR. NORBERT PESHU,**
THE DIRECTOR, CGMR-C,
KILIFI

forwarded 12/9/10

RE: SSC PROTOCOL NO. 1651 (*REQUEST FOR ANNUAL RENEWAL*): HOUSEHOLD TRANSMISSION OF RESPIRATORY SYNCYTIAL VIRUS (RSV): WHO ACQUIRES INFECTION FROM WHOM?

This is to inform you that during the 182nd meeting of the KEMRI/ERC meeting held on 6th September 2010, the Committee conducted the annual review and approved the above referenced application for another year.

This approval is valid from today September 10, 2010 through to September 9, 2011. Please note that authorization to conduct this study will automatically expire on September 9, 2011. If you plan to continue with data collection or analysis beyond this date please submit an application for continuing approval to the ERC Secretariat by August 5, 2011.

Yours sincerely,

R. C. KITHINJI,
FOR: SECRETARY,
KEMRI/NATIONAL ETHICS REVIEW COMMITTEE

KEMRI / WELLCOME TRUST RESEARCH PROGRAMME
17 SEP 2010
RECEIVED
P. O. Box 230-80108, KILIFI

c) Second annual approval letter

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840 - 00200 NAIROBI, Kenya
Tel: (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **September 14, 2011**

TO: **PATRICK MUNYWOKI (PRINCIPAL INVESTIGATOR)**

Forwarded: 20/09/11

THROUGH: **DIRECTOR, CGMR-C**
KILIFI

DIRECTOR
CENTRE FOR GEOGRAPHIC MEDICINE RESEARCH, COAST

Dear Sir,

RE: **SSC PROTOCOL No. 1651(*REQUEST FOR ANNUAL RENEWAL*): HOUSEHOLD TRANSMISSION OF RESPIRATORY SYNCYTIAL VIRUS (RSV): WHO ACQUIRES INFECTION FROM WHOM?**

Thank you for the Continuing Review Report for the period 10th September 2010 to 9th September, 2011.

This is to inform you that at the 193rd meeting of the KEMRI Ethics Review Committee held on 13th September 2011, your request for continuing approval was considered.

The Committee noted that:

a. Field work was completed on 31st June 2010.
b. 60 households were enrolled with a total of 596 participants. 13 household withdrew. On average each household had 10 participants with a median (interquartile range, IQR) of 8 (6, 11).
c. A total of 19,816 home visits resulting in 16,284 (82% nasal swabs and 9,226 (93%) oral fluids (OF) from 554 participants collected. In 3,624 (18%) visits the participants had symptoms of acute respiratory infection (ARI)
d. The planned activities for the next project year include screening of nasal swabs using multiplex PCR for 7 respiratory virus, sequencing of RSV and rhinovirus positive samples; finalising of data cleaning and descriptive analysis and preparing the data for inferential analysis once the nasal sample screening is completed.

The Committee concluded that the progress made in the reporting period is satisfactory and grants approval for continuation with the study effective **September 14, 2011**. Please note that authorization to conduct this study will automatically expire on **September 12, 2012**.

If you plan to continue with data collection or analysis beyond this date, please submit an application for continuing approval to the ERC Secretariat by **1st August, 2012**.

KEMRI/WELLCOME TRUST RESEARCH PROGRAMME
20 SEP 2011
RECEIVED
P. O. Box 230-80108 KILIFI

d) Third annual approval letter

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **September 26, 2012**

TO: **MR. PATRICK MUNYWOKI (PRINCIPAL INVESTIGATOR)**

THROUGH: **DR. SABAH OMAR,**
THE DIRECTOR, CGMRC,
KILIFI

forwarded 28/9/12

Dear Sir,

RE: **SSC PROTOCOL No. 1651 (*RE-SUBMISSION: REQUEST FOR ANNUAL RENEWAL*): HOUSEHOLD TRANSMISSION OF RESPIRATORY SYNCYTIAL VIRUS (RSV): WHO ACQUIRES INFECTION FROM WHOM?**

Reference is made to your letter dated September 19, 2012. The ERC Secretariat acknowledges receipt of the revised Continuing Review Report of September 20, 2012.

This is to inform you that the Committee notes that the reporting period is defined as 1st August 2011 to 18th July 2012 and is of the view that the progress made in the reporting period is satisfactory. Consequently, the study is granted approval for continuation effective this **26th day of September 2012.** Please note that authorization to conduct this study will automatically expire on **September 25, 2013.** If you plan to continue with data collection or analysis beyond this date, please submit an application for continuing approval to the ERC Secretariat by **August 14, 2013.**

You are required to submit any proposed changes to this study to the SSC and ERC for review and the changes should not be initiated until written approval from the ERC is received.

Please note that any unanticipated problems resulting from the conduct of this study should be brought to the attention of the ERC and you should advise the ERC when the study is completed or discontinued.

You may continue with the study.

DR. CHRISTINE WASUNNA,
ACTING SECRETARY,
KEMRI ETHICS REVIEW COMMITTEE

KEMRI / WELLCOME TRUST RESEARCH PROGRAMME
28 SEP 2012
RECEIVED
P O Box [illegible]

Appendix Q. Sensitivity of Qiagen and HP extraction methods in detection of RSV A, B and adenovirus

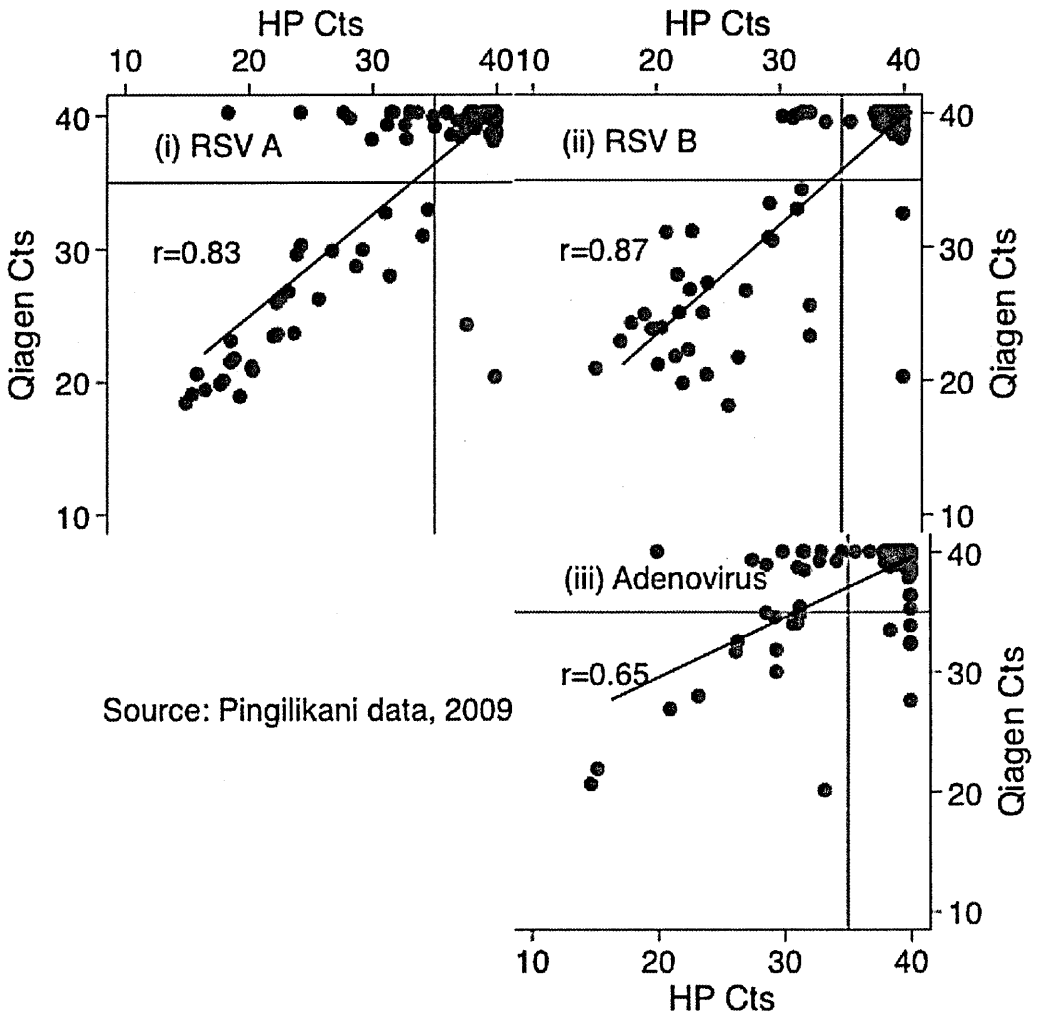
Samples used in the experiment reported in chapter3 were used. The 111 nasal samples had been screened by the 'RSV A and B and adenovirus' triplex assay based on Qiagen extraction method. Comparing the Qiagen results with the HP method allowed for assessing the differences in the two extraction methods, Table. The HP method had higher sensitivity in detection of RSV A and B relative to Qiagen method which was statistically significant for RSV A (p=0.0074). On comparison of the Ct values, HP had statistically significant lower Cts than Qiagen for all the viruses except for RSV B which was marginal (p=0.062). The correlation of the Ct values by the two methods was 83%, 87% and 65% for RSV A, B and adenovirus respectively as shown in Figure below.

I. Table: Respiratory pathogens detected and their sensitivity by extraction method

Viruses			HP Sensitivity		Qiagen Sensitivity	
	N	n	% (95% CI)	n	% (95% CI)	P value[1]
Any Virus	99	86	86.9 (78.6 – 92.8)	81	81.2 (72.79 – 88.9)	0.3593
RSV A	45	41	91.1(78.8 – 97.5)	30	66.7 (51.0 – 80.0)	0.0074
RSV B	37	35	94.6 (81.8 - 99.3)	30	81.1 (64.8 – 92.0)	0.1797
Adenovirus	39	22	56.4 (39.6 – 72.2)	28	71.8 (55.1 – 85.0)	0.2632

Key: N, total number of virus detections by PCR using either HP or TNA RNA extraction kits which is used as the reference for calculation of sensitivity; n, number of positive samples by the respective kit; %, sensitivity; CI, confidence interval; RSV, respiratory syncytial virus; 1, exact McNemar's significance probability values reported.

II. Figure showing the correlation of Qiagen with HP triplex Ct Values for RSV A, RSV B and Adenovirus.



Source: Pingilikani data, 2009

Appendix R. Distribution of Ct values

a) Approach

For this exploratory analysis, the frequency distributions of the Ct values for the various targets in the M-PCR assay were examined. The mean Ct value for the different viruses by age was also assessed. The aim was to check for any age-related association of the recorded Ct values.

b) Findings

Of the 16924 screened NPS collections, 1226 (7.2%) recorded a Ct value (value of greater than zero) for either RSV group A or B. Of the common respiratory pathogens screened for, a bimodal distribution was observed in RSV group A and B, NL63 and 2293 Ct values (Figure Ax). Adenovirus, OC43 and rhinovirus had a unimodal distribution of the recorded Ct values. The mean Ct values increased with age of the individual from whom the NPS was collected. Individuals aged 5 years or more had higher mean Ct values than the younger counterparts. This phenomenon was observed for the other targets (adenovirus and coronaviruses) except human rhinovirus where older ages (>15years) had similar Ct values as the under one year olds.

Table: Mean Ct values stratified by age for the most prevalent viruses as detected by M-PCR

Virus	Age, years	n	Mean Ct	95% CI	P-value
RSV[1]	<1y	232	30.4	29.6 – 31.1	Ref
	1-<5y	253	31.8	30.1 – 33.6	0.005
	5-<15	413	33.5	31.9 – 35.2	<0.0001
	15- <40y	241	35.0	33.2 – 36.8	<0.0001
	≥40	87	35.0	32.8 – 37.1	<0.0001
Adenovirus	<1y	237	31.4	30.9 – 32.0	Ref
	1-<5y	507	31.2	30.0 – 32.3	0.418
	5-<15	678	32.8	31.7 – 34.0	<0.0001
	15- <40y	255	33.7	32.5 – 35.0	<0.0001
	≥40	71	33.9	32.3 – 35.6	<0.0001
Coronavirus[2]	<1y	249	25.9	24.9 – 26.8	Ref
	1-<5y	369	27.5	25.3 – 29.7	0.009
	5-<15	604	28.7	26.5 – 30.8	<0.0001
	15- <40y	340	30.1	27.9 – 32.4	<0.0001
	≥40	90	29.3	26.4 – 32.1	<0.0001
Rhinovirus	<1y	511	30.2	29.8 – 30.6	Ref

1-<5y	521	31.3	30.4 – 32.2	<0.0001
5-<15	774	31.3	30.4 – 32.1	<0.0001
15- <40y	259	30.5	29.5 – 31.6	0.322
≥40	64	31.0	29.5 – 32.5	0.162

Key: 1, comprise of the best (lowest) Ct values for RSV group A or B; 2, comprise of the best (lowest) Ct values for OC43, NL63 and 229E strains; RSV, respiratory syncytial virus; Ref, reference group

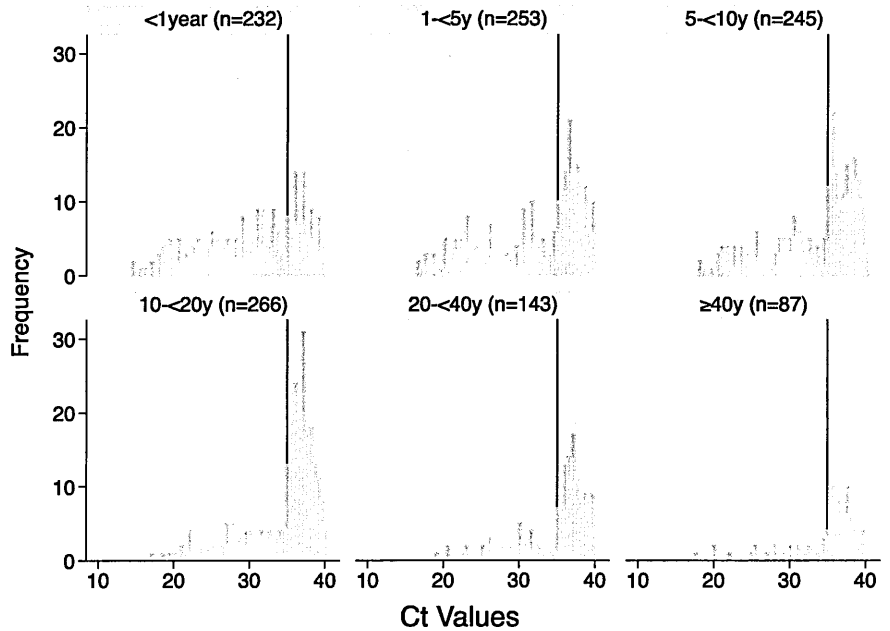


Figure A1: Frequency distribution of Ct values for RSV (group A or B) detection by age groups in years

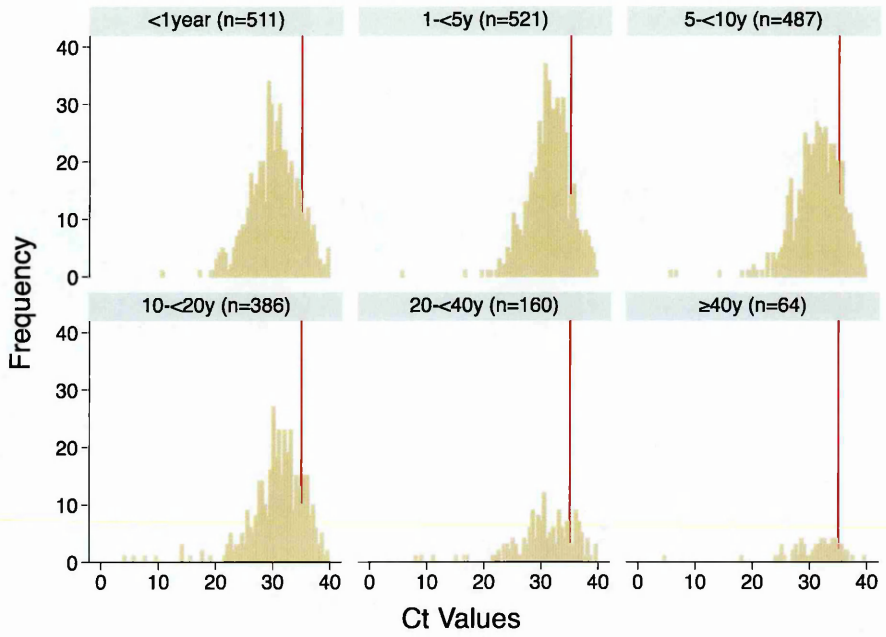


Figure A2: Frequency distribution of Ct values for human rhinovirus detection by age groups in years

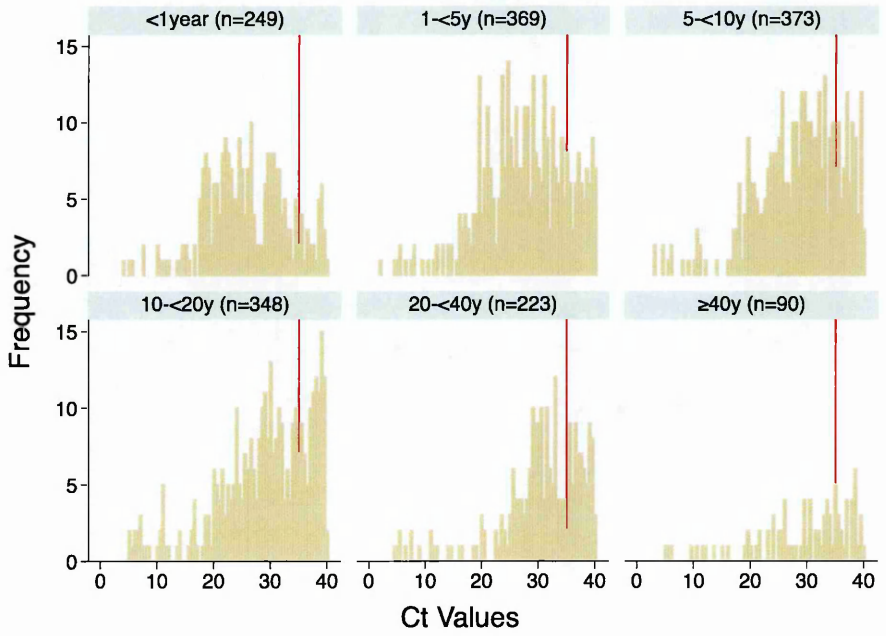


Figure A3: Frequency distribution of Ct values for human coronaviruses (OC42, NL63 or 229E) detection by age groups in years

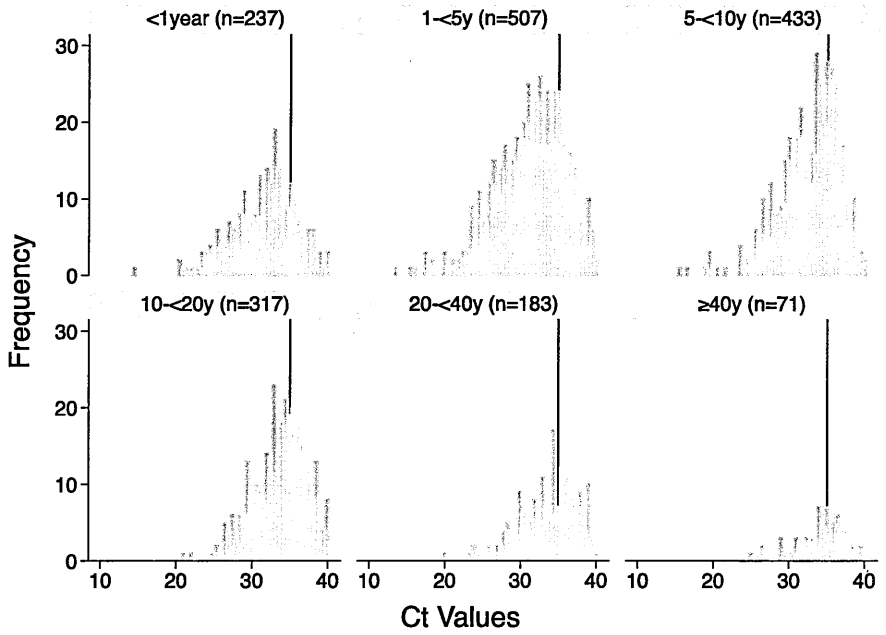


Figure A4: Frequency distribution of Ct values for adenovirus detection by age groups in years

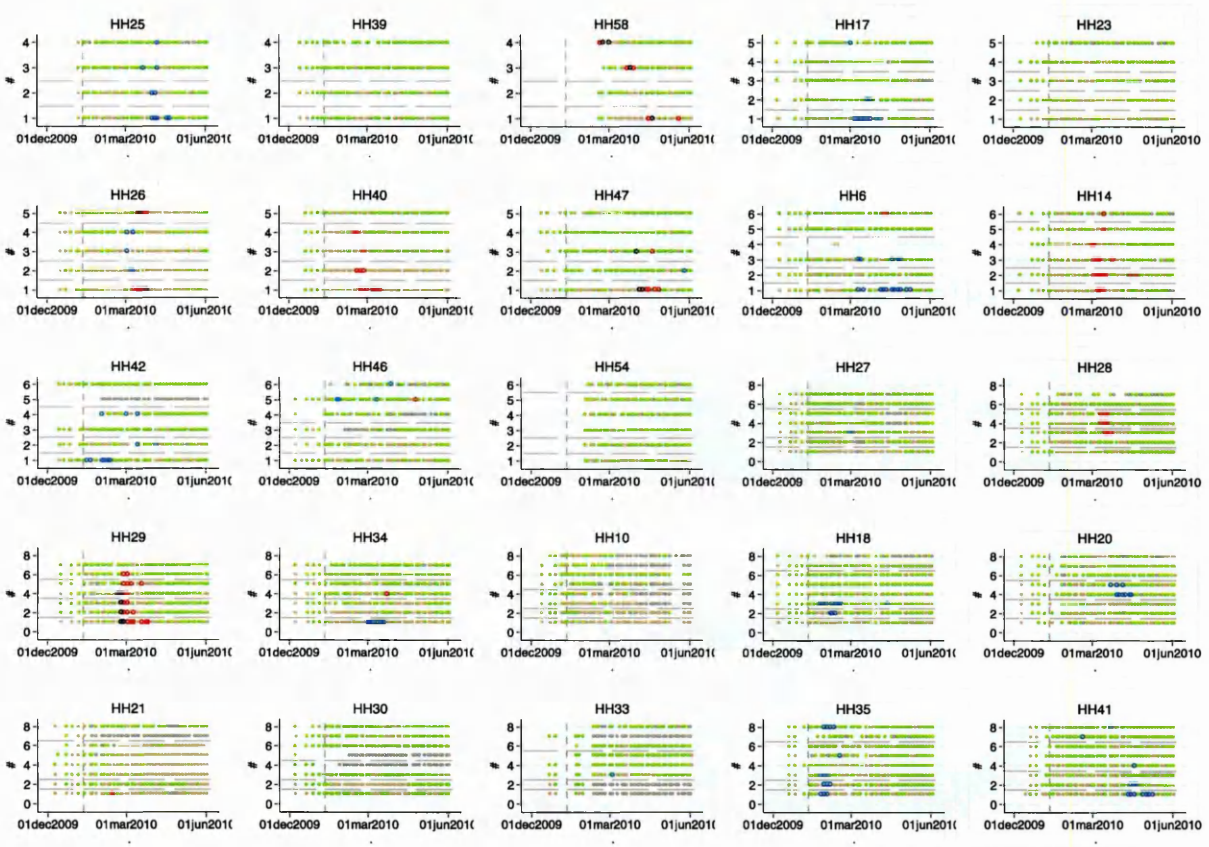
Appendix S. Sensitivity of oral fluid in detection of RSV by Fast Track diagnostics

Characteristic	Categories	n	No. of RSV Positives	Sensitivity (%)	95% CI
Overall	All	58	16	27.6	16.7 – 40.9
NPS CTs	31 – 35	23	4	17.4	5.0 – 38.8
	26 – 30	25	7	28.0	12.1 – 49.4
	≤25	16	6	37.5	15.2 – 64.6
RSV Group	RSV A	36	12	33.3	18.6 – 51.0
	RSV B	28	5	17.4	6.1 – 36.9
Age (years)	<5 y	26	12	46.2	26.7 – 66.6
	≥5 y	33	4	12.1	3.4 – 28.2
Age groups	<1 y	9	4	44.4	13.7 – 78.8
	1 – <5 y	17	8	47.1	23.0 – 72.2
	5 – <10 y	13	2	15.3	1.9 – 45.4
	≥10 y	20	2	10.0	1.2 – 31.7
Gender	Female	29	9	31.0	15.3 – 50.8
	Male	30	7	23.3	9.9 – 42.3
ARI	Yes	23	8	34.8	16.4 – 57.3
	No	36	8	22.2	10.1 – 39.2

For the Fast Track Diagnostics, the RNA in the samples was extracted using the EasyMAG NucliSENS Extractor and tested for RSV by fast-track diagnostics respiratory pathogens multiplex real-time RT-PCR assay following manufacturers instructions and as described elsewhere (Sakthivel et al. 2012)

Appendix T. RSV infection patterns in the 47 study households

I. NPS collection and RSV infection patterns in the 25 households with 8 members



II. NPS collection and RSV infection patterns in 22 households with >8 members

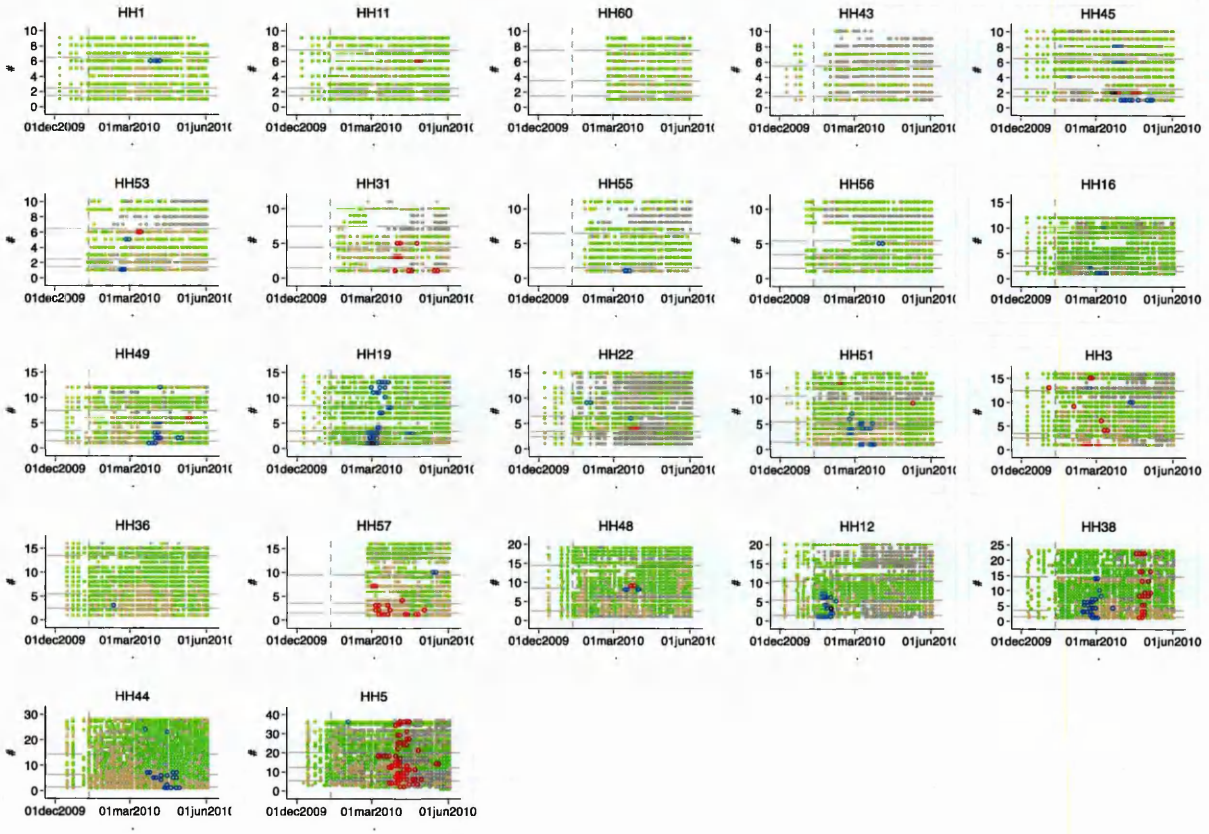


Figure legend: Each panel represents a household with the subtitle showing the household ID, sorted by household size. The y-axis shows the household members, sorted by age with the youngest at number 1. The dashed grey (horizontal) lines show the cut off for 1 year, 5 years and 15 years of ages, at start of sampling, in each household. The sample collection dates are shown on the x-axis. The dashed light grey vertical line shows the start of the main study phase. The green circles indicate PCR negative NPS collections while the PCR positives are shown in red, blue and purple for RSV group A, group B and both (group A and B), respectively. The maroon filled circles show visits when the participant had ARI symptoms while the grey empty circles indicate when the participants were away from the household.

Appendix U. Figure: RSV shedding patterns for the 205 episodes

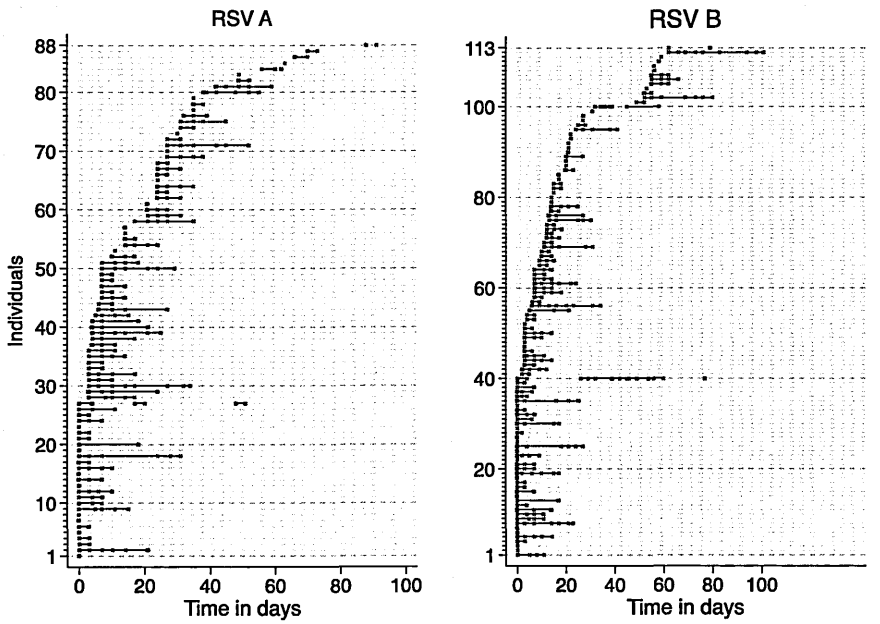
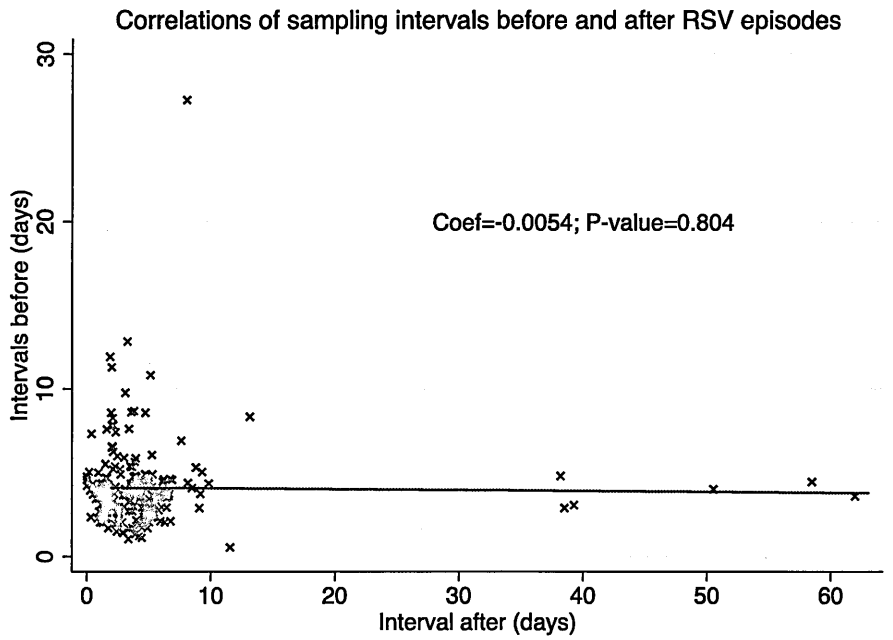


Figure: RSV infection episodes by order of occurrence within their respective households.

Appendix V. Figure: The correlation of sampling intervals before and after RSV episodes



Appendix W. Table: Mean duration of the 205 RSV infection episodes by estimation method among rural Kenyan children

Factor	Categories	Recover[1]	C[2]	Minimum estimates			Midpoint estimates			Maximum estimates		
				Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
Overall		190	15	8.60	7.50	9.70	11.20	10.06	12.30	14.00	12.80	15.20
Either left or right	No	177	0	8.23	7.07	9.39	10.92	9.75	12.10	13.62	12.43	14.80
Censored	Yes	13	15	10.05	7.40	12.69	12.22	9.60	14.85	17.90	12.00	23.79
Left censored episodes	No	179	15	8.73	7.54	9.92	11.43	10.24	12.61	14.24	13.01	15.47
	Yes	11	0	6.68	4.04	9.32	7.37	4.65	10.08	9.91	7.06	12.76
Age groups in years	<1y	30	3	15.47	11.80	19.13	17.97	14.17	21.77	20.61	16.69	24.53
	1-4y	47	5	9.24	7.10	11.37	11.84	9.66	14.02	14.55	12.31	16.78
	5-14y	72	3	6.42	5.04	7.80	9.07	7.71	10.44	11.86	10.47	13.25
	15-39y	34	4	5.53	3.72	7.34	8.39	6.59	10.19	11.60	9.30	13.91
	≥40y	7	2	9.43	5.54	13.32	11.19	7.38	15.00	13.51	9.72	17.30
Relationships	Infants	34	0	14.73	11.32	18.13	17.26	13.71	20.80	19.90	16.23	23.56
	Siblings	79	0	6.86	5.42	8.31	9.36	7.91	10.82	12.00	10.53	13.48
	Cousins	45	10	8.32	6.42	10.21	11.33	9.47	13.19	14.66	12.49	16.83

	Mother	14	0	6.49	3.58	9.40	8.92	6.10	11.74	11.48	8.64	14.32
	Father	4	2	4.93	2.57	7.29	6.83	4.10	9.57	9.54	6.67	12.40
	Others	14	3	6.59	3.24	9.93	9.15	5.97	12.34	11.64	8.61	14.67
Gender	Female	107	9	8.73	7.06	10.39	11.22	9.54	12.89	13.93	12.22	15.64
	Male	83	6	8.49	7.01	9.96	11.24	9.75	12.72	14.14	12.55	15.73
RSV group	A	81	12	9.52	7.85	11.20	11.95	10.29	13.62	14.69	13.00	16.38
	B	109	3	7.85	6.35	9.34	10.57	9.04	12.09	13.42	11.82	15.02
RSV infection	First	164	15	8.95	7.73	10.16	11.47	10.24	12.70	14.21	12.97	15.46
	Second	26	0	6.31	3.51	9.10	9.38	6.52	12.25	12.46	9.10	15.83
Symptomatic episode	No	78	9	5.25	4.15	6.35	7.76	6.70	8.81	10.53	9.45	11.61
	Yes	112	6	10.86	9.27	12.45	13.52	11.90	15.14	16.33	14.64	18.02
Others infected in household (%)	0 – <33%	76	3	5.82	4.52	7.12	8.67	7.32	10.02	11.71	10.15	13.27
	33 – <66%	66	1	9.75	7.60	11.90	12.21	10.02	14.39	14.76	12.55	16.97
	≥ 0.66%	48	11	11.00	8.81	13.20	13.48	11.28	15.67	16.20	14.00	18.39
During household outbreak	No	39	3	4.68	3.21	6.16	7.30	5.82	8.78	10.05	8.49	11.61
	Yes	151	12	9.58	8.26	10.90	12.18	10.85	13.50	14.99	13.61	16.36

Primary cases of household episodes	No	110	11	8.94	7.49	10.40	11.43	9.96	12.90	14.11	12.61	15.61
	Yes	80	4	8.08	6.30	9.86	10.85	9.07	12.63	13.80	11.92	15.68
Primary cases of household outbreaks	No	151	15	7.84	6.66	9.01	10.43	9.25	11.62	13.23	11.98	14.48
	Yes	39	0	11.60	8.73	14.47	14.20	11.32	17.08	17.00	14.10	19.89
Presence of other virus (co-infections)	No	110	9	5.85	4.75	6.95	8.45	7.39	9.52	11.32	10.15	12.49
	Yes	80	6	12.30	10.37	14.23	14.84	12.86	16.82	17.55	15.53	19.57
Other viruses detected during RSV episode	None	110	9	5.85	4.75	6.95	8.45	7.39	9.52	11.32	10.15	12.49
	Adenoviruses	21	1	8.83	5.79	11.87	11.63	8.47	14.79	14.47	11.14	17.80
	Coronaviruses	20	2	12.53	8.87	16.20	15.04	11.39	18.69	17.70	14.10	21.31
	Rhinoviruses	19	0	14.30	9.63	18.98	16.64	11.75	21.53	19.18	14.11	24.25
	Mixed	20	3	13.50	9.99	17.00	16.05	12.43	19.67	18.90	15.22	22.57
No. of the co-infecting viruses during RSV episodes	0	110	9	5.85	4.75	6.95	8.45	7.39	9.52	11.32	10.15	12.49
	1	60	3	11.83	9.55	14.11	14.40	12.07	16.74	17.07	14.69	19.45
	≥2	20	3	13.50	9.99	17.00	16.05	12.43	19.67	18.90	15.22	22.57
Other viruses 14 days prior to RSV episode	No	92	9	7.91	6.41	9.41	10.55	9.06	12.05	13.49	11.90	15.08
	Yes	98	6	9.26	7.59	10.92	11.79	10.10	13.47	14.44	12.73	16.16

Other viruses detected prior to RSV infection	None	92	9	7.91	6.41	9.41	10.55	9.06	12.05	13.49	11.90	15.08
	Adenovirus	20	1	8.23	5.27	11.19	10.71	7.78	13.64	13.36	10.35	16.37
	Coronavirus	18	2	8.02	4.88	11.16	10.48	7.40	13.56	13.27	10.23	16.31
	Rhinovirus	36	1	8.97	5.93	12.01	11.58	8.49	14.67	14.19	11.06	17.33
	Mixed	24	2	11.42	7.95	14.88	13.91	10.32	17.49	16.54	12.85	20.23
No. of other viruses detected 14 days prior to RSV episodes	0	92	9	7.91	6.41	9.41	10.55	9.06	12.05	13.49	11.90	15.08
	1	74	4	8.57	6.69	10.44	11.11	9.23	12.99	13.77	11.86	15.67
	≥2	24	2	10.82	7.24	14.40	13.27	9.57	16.97	15.88	12.07	19.69
Other viruses detected 14 days prior to and/or during RSV episode	None	69	6	6.38	4.89	7.86	8.97	7.58	10.36	12.00	10.42	13.58
	Before only	41	3	5.17	3.43	6.91	7.57	5.94	9.20	10.14	8.55	11.73
	Before & during	57	3	12.16	9.88	14.45	14.74	12.40	17.08	17.45	15.05	19.85
	During only	23	3	12.71	9.11	16.31	15.22	11.53	18.90	17.86	14.14	21.58

Key: 1, episodes with observed recovery time; 2, right-censored RSV infection episodes; LCL, lower confidence limit; UCL, upper confidence limit

Appendix X. Table: Recovery rates of RSV infection among rural Kenyan children by the three estimation methods

Factor	Categories	Minimum estimates			Midpoint estimates			Maximum estimates		
		Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
Overall		11.59	10.05	13.36	8.73	7.57	10.06	6.92	6.00	7.97
Either left or right Censored	No	7.11	4.13	12.24	5.34	3.10	9.20	3.86	2.24	6.65
	Yes	12.15	10.48	14.08	9.15	7.90	10.61	7.34	6.34	8.51
Left censored episode	No	11.43	9.87	13.23	8.54	7.38	9.89	6.79	5.86	7.86
	Yes	14.96	8.29	27.02	13.58	7.52	24.52	10.09	5.59	18.22
Age groups in years at infection	<1y	6.30	4.40	9.01	5.40	3.78	7.73	4.68	3.28	6.70
	1-4y	10.70	8.04	14.24	8.19	6.15	10.89	6.57	4.94	8.74
	5-14y	15.55	12.34	19.59	10.93	8.68	13.77	8.35	6.63	10.52
	15-39y	18.37	13.13	25.72	11.45	8.18	16.02	8.17	5.84	11.44
	>40y	9.16	4.37	19.22	7.62	3.63	15.98	6.21	2.96	13.02
Relationships	Study infants	6.79	4.85	9.50	5.79	4.14	8.11	5.03	3.59	7.03
	Siblings	14.57	11.69	18.17	10.68	8.57	13.32	8.33	6.68	10.39
	Cousins	11.62	8.68	15.57	8.14	6.07	10.90	6.17	4.61	8.26

	Mother	15.41	9.13	26.02	11.21	6.64	18.93	8.71	5.16	14.71
	Father	17.76	6.67	47.32	11.76	4.41	31.33	8.16	3.06	21.73
	Others	14.51	8.59	24.50	10.11	5.99	17.07	7.65	4.53	12.92
Gender	Females	11.48	9.50	13.87	8.75	7.24	10.57	6.96	5.76	8.42
	Males	11.73	9.46	14.55	8.70	7.02	10.79	6.86	5.53	8.51
RSV group	A	10.33	8.31	12.84	7.96	6.40	9.90	6.38	5.13	7.93
	B	12.74	10.56	15.37	9.40	7.79	11.34	7.38	6.12	8.90
Episode number	First	11.11	9.53	12.95	8.48	7.28	9.89	6.77	5.81	7.89
	Second	15.85	10.79	23.28	10.66	7.26	15.65	8.02	5.46	11.79
Symptomatic episodes	No	18.99	15.21	23.71	12.28	9.83	15.33	8.90	7.13	11.11
	Yes	9.11	7.57	10.97	7.26	6.04	8.74	5.99	4.98	7.21
Presence of the other infected members in household (%)	0 – <33%	17.32	13.83	21.69	11.43	9.13	14.31	8.42	6.72	10.54
	33 – <66%	10.24	8.05	13.04	8.17	6.42	10.39	6.73	5.29	8.57
	≥ 0.66%	8.62	6.50	11.44	6.82	5.14	9.05	5.56	4.19	7.37
Episode during a household outbreak	No	21.75	15.89	29.76	13.41	9.80	18.35	9.59	7.01	13.12
	Yes	10.34	8.81	12.13	8.01	6.83	9.39	6.45	5.50	7.57

Primary cases for household episodes	No	11.03	9.15	13.30	8.46	7.02	10.20	6.78	5.62	8.17
	Yes	12.45	10.00	15.50	9.12	7.33	11.36	7.12	5.72	8.86
Primary case for an outbreak	No	12.72	10.84	14.91	9.30	7.93	10.91	7.25	6.18	8.50
	Yes	8.62	6.30	11.80	7.04	5.14	9.64	5.88	4.30	8.05
Presence of co-infections during RSV episode	No	17.20	14.27	20.73	11.47	9.52	13.83	8.49	7.04	10.23
	Yes	8.00	6.42	9.96	6.57	5.27	8.18	5.52	4.43	6.87
Co-infections	None	17.20	14.27	20.73	11.47	9.52	13.83	8.49	7.04	10.23
	Adenoviruses	11.18	7.29	17.15	8.46	5.51	12.97	6.78	4.42	10.40
	Coronaviruses	7.80	5.03	12.08	6.41	4.14	9.94	5.39	3.48	8.35
	Rhinoviruses	6.99	4.46	10.96	6.01	3.83	9.42	5.21	3.33	8.17
	Mixed	7.04	4.54	10.91	5.85	3.78	9.07	4.93	3.18	7.65
No. of different co-infecting viruses	0	17.20	14.27	20.73	11.47	9.52	13.83	8.49	7.04	10.23
	1	8.38	6.51	10.79	6.84	5.31	8.82	5.74	4.46	7.39
	2	7.04	4.54	10.91	5.85	3.78	9.07	4.93	3.18	7.65
Other viruses detected 14 days prior to RSV episode	No	12.60	10.27	15.45	9.17	7.48	11.25	7.09	5.78	8.70
	Yes	10.77	8.84	13.13	8.35	6.85	10.17	6.76	5.55	8.24

Viruses detected 14 days prior to RSV infection	None	12.60	10.27	15.45	9.17	7.48	11.25	7.09	5.78	8.70
	Adenoviruses	11.92	7.69	18.47	9.08	5.86	14.07	7.31	4.72	11.33
	Coronaviruses	12.45	7.85	19.77	9.18	5.79	14.57	7.17	4.52	11.38
	Rhinoviruses	11.19	8.07	15.51	8.62	6.22	11.95	6.97	5.03	9.66
	Mixed	8.71	5.84	13.00	7.06	4.73	10.53	5.88	3.94	8.78
No. of different viruses 14 days prior to start of RSV episode	0	12.60	10.27	15.45	9.17	7.48	11.25	7.09	5.78	8.70
	1	11.67	9.29	14.66	8.87	7.06	11.14	7.11	5.66	8.93
	2	9.25	6.15	13.92	7.46	4.96	11.23	6.21	4.13	9.34
	3	3.73	0.53	26.46	3.14	0.44	22.30	2.66	0.37	18.85
Viruses detected 14 days prior to and during RSV episode	None	15.96	12.61	20.21	10.76	8.50	13.62	7.98	6.30	10.10
	Before only	19.77	14.56	26.85	12.92	9.51	17.55	9.50	6.99	12.90
	Before & during	8.12	6.26	10.52	6.65	5.13	8.63	5.60	4.32	7.26
	During only	7.72	5.13	11.61	6.36	4.23	9.57	5.31	3.53	7.99

Key: 1 – number of recovered infections; 2 number of right-censored infections; 3 total person-days of follow up; 4 rate of recovery; 5 lower 95% confidence Limit; 6 upper 95% confidence limit

Appendix Y. Table: Unadjusted hazard ratios of RSV recovery for various covariates from univariate Cox regression analysis

Factors		Reference	Minimum estimates				Midpoint estimates				Maximum estimates			
			HR	LCL	UCL	P-value	HR	LCL	UCL	P-value	HR	LCL	UCL	P-value
Age groups	1-4y	<1year	1.89	1.17	3.06	0.009	1.91	1.18	3.09	0.008	1.78	1.10	2.86	0.018
	5-14y		2.75	1.73	4.37	<0.001	2.78	1.75	4.42	<0.001	2.53	1.61	3.98	<0.001
	15-39y		3.19	1.88	5.42	<0.001	3.23	1.90	5.48	<0.001	2.61	1.56	4.36	<0.001
	≥40y		1.80	0.77	4.20	0.173	1.92	0.83	4.49	0.13	1.85	0.80	4.29	0.152
Relationships	Siblings*	Study infants	2.14	1.42	3.21	<0.001	2.19	1.46	3.29	<0.001	2.02	1.35	3.01	0.001
	Others		2.46	1.47	4.11	0.001	2.70	1.61	4.51	<0.001	2.57	1.54	4.28	<0.001
Symptomatic episode		Asymptomatic	0.48	0.35	0.65	<0.001	0.45	0.33	0.61	<0.001	0.47	0.34	0.64	<0.001
Presence of other viruses		None	0.45	0.33	0.61	<0.001	0.43	0.32	0.59	<0.001	0.47	0.35	0.64	<0.001
Other viruses detected	Adeno	None	0.66	0.41	1.05	0.08	0.61	0.38	0.97	0.038	0.63	0.40	1.02	0.059
	Corona		0.44	0.27	0.72	0.001	0.42	0.26	0.69	0.001	0.47	0.29	0.77	0.003
	Rhino		0.36	0.22	0.61	<0.001	0.35	0.21	0.59	<0.001	0.40	0.24	0.67	<0.001
	Mixed		0.40	0.25	0.66	<0.001	0.38	0.23	0.62	<0.001	0.42	0.26	0.68	<0.001
Others	33 – <66%	<33%	0.59	0.42	0.83	0.003	0.61	0.43	0.86	0.005	0.68	0.48	0.95	0.024

infected in HH (%)	≥66%		0.53	0.36	0.76	0.001	0.54	0.38	0.78	0.001	0.58	0.41	0.84	0.004
Primary cases of HH episodes		No	1.11	0.83	1.49	0.472	1.08	0.81	1.44	0.614	1.03	0.77	1.38	0.822
Primary cases of HH outbreaks		No	0.69	0.48	0.99	0.043	0.68	0.47	0.97	0.033	0.69	0.48	0.99	0.044
During HH outbreak		No	0.51	0.35	0.73	<0.001	0.49	0.34	0.71	<0.001	0.52	0.36	0.74	<0.001
No. of other viruses	1	None	0.47	0.34	0.66	<0.001	0.45	0.32	0.62	<0.001	0.49	0.35	0.68	<0.001
	2		0.40	0.24	0.66	<0.001	0.38	0.23	0.63	<0.001	0.41	0.25	0.69	0.001
	≥3		0.50	0.12	2.05	0.339	0.47	0.12	1.90	0.288	0.48	0.12	1.97	0.31
Male gender		Female	1.03	0.77	1.38	0.836	1.02	0.76	1.36	0.916	0.99	0.74	1.32	0.919
RSV group	B	A	1.08	0.79	1.46	0.642	1.04	0.77	1.42	0.782	1.03	0.76	1.40	0.852
	Both		0.53	0.29	0.97	0.039	0.52	0.28	0.94	0.032	0.55	0.30	1.00	0.048
Second RSV infection		First	1.36	0.90	2.07	0.149	1.28	0.84	1.94	0.248	1.20	0.79	1.82	0.4
Other viruses detected prior to RSV infection		None	0.85	0.64	1.14	0.283	0.86	0.65	1.15	0.32	0.90	0.67	1.20	0.463
Viruses	Adeno	No detections	0.97	0.60	1.58	0.914	1.00	0.61	1.62	0.984	1.03	0.63	1.67	0.914

detected prior to RSV infection	Corona		1.00	0.60	1.66	0.996	1.01	0.61	1.67	0.977	1.05	0.63	1.74	0.862
	Rhino		0.86	0.58	1.27	0.445	0.87	0.58	1.28	0.472	0.91	0.61	1.35	0.635
	Mixed		0.70	0.44	1.09	0.117	0.70	0.45	1.11	0.129	0.73	0.46	1.14	0.167
No. of viruses detected prior to RSV infection	1	0	0.92	0.68	1.25	0.603	0.93	0.68	1.27	0.655	0.97	0.71	1.32	0.854
	2		0.74	0.47	1.17	0.195	0.75	0.47	1.19	0.222	0.77	0.49	1.22	0.272
	≥3		0.31	0.04	2.22	0.244	0.30	0.04	2.13	0.226	0.31	0.04	2.23	0.245
Other viruses detected prior to and during RSV episode	Before only	None	1.21	0.82	1.79	0.334	1.31	0.89	1.94	0.172	1.39	0.94	2.06	0.096
	Both		0.49	0.34	0.70	<0.001	0.48	0.33	0.69	<0.001	0.53	0.37	0.76	0.001
	During only		0.47	0.29	0.77	0.002	0.46	0.29	0.75	0.002	0.52	0.32	0.84	0.007
Student		Non-students	1.28	0.94	1.73	0.112	1.22	0.90	1.65	0.206	1.12	0.83	1.52	0.457
Left censor		Not left censor	1.55	0.84	2.86	0.163	2.11	1.14	3.91	0.017	2.19	1.18	4.04	0.013

Key: HR, hazard ratio; LCL, lower confidence limit; UCL, upper confidence limit; adeno, adenoviruses; corona, coronaviruses; rhino, rhinoviruses

Appendix Z. Table: Multivariate Cox regression model including left censored RSV episodes

Factors	Categories	Reference	HR	95% CI	p-value
Age groups	1-4y	<1 year	1.85	1.22 – 2.80	0.004
	5-14y		1.77	1.14 – 2.75	0.01
	≥15y		1.67	0.93 – 2.99	0.084
Symptomatic		Asymptomatic	0.59	0.42 – 0.82	0.002
Detection of other[1] viruses before and during RSV episode	Before[2]	No other viruses	1.55	1.02 – 2.35	0.041
	During		0.50	0.35 – 0.70	<0.001
Other HH members infected (%)	33 – <66%	<33%	0.58	0.42 – 0.82	0.002
	≥ 66%		0.57	0.41 – 0.80	0.001
RSV group B		A	1.24	0.96 – 1.61	0.1
Male gender		Female	1.00	0.78 – 1.27	0.992
Second RSV infection		First	0.95	0.54 – 1.68	0.857
Left censoring		Not left censored	1.88	0.76 – 4.65	0.172

Key: HR, hazard ratio; CI, confidence interval; HH, household; 1, detection of other viruses (rhinoviruses, adenoviruses and coronaviruses) during the RSV episode; 2, detection of other respiratory viruses during the 14 days prior to the start of RSV episode ONLY

Appendix AA. Table: Test of proportional-hazards assumption using Schoenfeld residuals

Factor	Categories	Rho	Chi-square statistic	P value
Age groups	1-4y	0.0028	0	0.9726
	5-14y	0.01222	0.02	0.8776
	15-39y	-0.01208	0.03	0.8597
Symptomatic		0.01549	0.05	0.8284
Detection of other viruses before and during RSV episode	Before	-0.00864	0.02	0.8989
	During	0.02406	0.12	0.7297
Other HH members infected (%)	33 – <66%	-0.03005	0.2	0.6517
	≥ 66%	0.02584	0.1	0.7461
RSV group B		0.00053	0	0.9949
Male gender		0.11266	1.62	0.203
Second RSV infection		-0.00577	0.01	0.9223
Global test			2.52	0.9957

Key: HH, household

Appendix BB. R code

I. Log likelihood function and fitting

```
#####################  START  ############################
# Written by PMunywoki April 2013,
# Modified by GFMedley May 2013 for faster execution
#Function for assigning proportionate contribution to infections by GFMedley May
2013
# Final code updated by PMunywoki in May 2013

#Working directory
setwd("/Users/pmunywoki/My Work/KILIFI/Household study/data/rstudio")
rm(list=ls())

#GET the data
source('create_data.R') # code shown below

#MANIPULATE the data
#make the contact parameter index array
#This array has values from 1..np where np is the number of contact #parameters+1.
#For example, if there are 3 contact parameters, then the values are 1,2,3,4 #where 2
indicates the 1st contact parameter. A value 1 indicates no HH #contact. The diagonal
is set to 1 to avoid self contact

cpa = rlns3gps
diag(cpa)<-0
cpa <- cpa+1

#Create the age matrix as an index for the age parameters
sus.A <- matrix(1,1,n)
sus.A[agegp1_5] <- 2 #the second parameter - the first parameter is a dummy,
                        so that the value is always 1
sus.A[agegp5_15] <- 3
sus.A[agegp15_40] <- 4
sus.A[agegp40] <- 5

#Create the post infection matrix
sus.B <- matrix(1,d,n) # post infection
sus.B [postInf] <- 2 #the second parameter

#Create the contact matrix as an index of the contact parameters
CPI <- rlns3gps + 1

#MAKE the likelihood functions
source("funDefs.R") # the script is defined below

#TESTING the function
pars <- cbind ( -4.605, -1.050, matrix(-1.5, 1, 4), matrix(-4, 1, 9), -0.916 )
```

```
ll <- cal.nLL(pars,sus, d,n,atRisk,shedDen,latency,onset.latency,CPI,HH8gps)
#llold <- cal.nLLold(pars,d,n,atRisk,shedDen,latency,onset.latency, postInf,rlns3gps,
HH8gps)

#OPTIMISE the function to get estimates of pars
A <- optim(pars,cal.nLL, gr=NULL, sus, d, n, atRisk, shedDen,
latency,onset.latency,
        CPI, HH8gps, method='BFGS',
         control = list(maxit = 500, trace = TRUE, REPORT = 5))

#Plot the age effects
agePars <- matrix ( c( 1, exp ( A$par[4:7]) ), 2, 5, byrow="TRUE" )
ageLims <- matrix ( c(0,1, 1,5, 5,15, 15,40, 40,100), 2, 5 )
plot(ageLims, agePars, type="o",main="Age parameters", xlab="Age (yrs)", ylab
="Susceptibility",ylim=c(0,1))

#Get the mixing parameters
conPars <- matrix ( c( 1, exp(A$par[8:15]) ), 3, 3 )
cat("Contact parameters:")
print(conPars)

#CALCULATE the individual contributions to infections
#Get the list of days and individuals when onset occurred
onsets <- which ( onset, arr.ind=TRUE)
colnames(onsets) <- c("DAY", "INDIV")

#Returns the relative contribution for each onset: 1...7 where 7 is the community
con <- cal.contrib (A$par, d, n, atRisk, shedDen,latency,onset.latency, postInf,
CPI, HH8gps,rlnship, onsets )

#Collect together the results for each infected relationship group
results = matrix ( 0, 6, 12 )
for ( ir in 1:6 ) {
 cc <-  con[rlnship[onsets[,2]]==ir,]
 results[ir,1] <- tabulate ( rlnship )[ir] #the number of individuals with this
                                           relationship status
 results[ir,3] <- dim(cc)[1] #the number of onsets with this relationship status
 results[ir,5:11] <- colSums(cc)
}
cat("Table of contributions:")
print(results)
```

II. Script For Data Input (Create_data.R)

```
#####create_data.R ##############
#Input data and create objects for use in the various models shown in the main text
#Working directory
setwd("/Users/pmunywoki/My Work/KILIFI/Household study/data/rstudio")
rm(list=ls())
```

```
#Load data
library(foreign)
hh.data <-  read.csv("hh_all.csv")
dim(hh.data)

##Create list of hh names
names <-substr(list.files(path="/Users/pmunywoki/My Work/KILIFI/Household
study/data/rstudio/hhs", pattern="hh_+.*csv"),1,5)

## RSV infections for all individuals
rsv <- as.matrix(hh.data)
rsv <- rsv[,-1]

#Number days of follow up
d <- as.numeric(nrow(hh.data))
days <- matrix(1:d,d,1)

#Number of participants
n <- as.numeric(ncol(rsv))

#Load age data
age.data <- as.matrix(read.csv('agey.csv')); sid <-  as.vector(age.data[,1]) ;

#agey
agey <- as.numeric(age.data[,2])

#Create age groups
#agegps <- c('agegp0','agegp1','agegp5','agegp15','agegp40')
agegp0_1 <- t(as.matrix(agey>=0 & agey<1)) ; colnames(agegp0_1) <- sid
agegp1_5 <- t(as.matrix(agey>=1 & agey<5)) ; colnames(agegp1_5) <- sid
agegp5_15 <- t(as.matrix(agey>=5 & agey<15)) ; colnames(agegp5_15) <- sid
agegp15_40 <- t(as.matrix(agey>=15 & agey<40)) ; colnames(agegp15_40) <- sid
agegp40 <- t(as.matrix(agey>=40)) ; colnames(agegp40) <- sid
agey <- t(as.matrix(as.numeric(agey))) ; colnames(agey) <- sid

#Create relationships
#rlns // 1 - self; 2 siblings/cousins; others
rlns <- t(age.data[,3]) ; colnames(rlns) <- sid
rlns.t <- t(rlns)

#rlns // 1 - self; 2 siblings; 3 cousins; 4 mother; 5 father; 6 others
rlnship <- as.numeric(age.data[,4]) ; rlnship <- t(as.matrix(rlnship))
      colnames(rlnship) <- sid
rlnship.t <- t(rlnship)

#Format dates
dates <- as.Date(c('01dec2009'), "%d%B%Y")
sampledates <- as.Date(seq(dates, (dates+185),1))
```

```
#HHID list
hh.list <- c(1,3,5,6,10,11,12,14,16,17,18,19,20,21,22,23,25,26,27,28,29,30,31,
        33,34,35,36,38,39,40,41,42,43,44,45,46,47,48,49,51,53,54,55,56,57,58,60)

#hh sizes by order of the hh
hh.sizes <- c(9,16,37,6,8,9,20,6,12,5,8,14,8,8,15,5,4,5,7,7,7,8,11,8,7,8,16,23,
         4,5,8,6,10,28,10,6,5,19,12,15,10,7,11,11,16,4,9)

#HHID matrix // both col and row matrix
HH.t <- as.matrix(as.numeric(substr(sid,4,5)))
HH <- t(HH.t)

#Household sizes
HHsizes <- matrix(0,1,n)
for (i in 1:47) {
 for (c in 1:n) {
  if (HH[c]==hh.list[i]) {HHsizes[c] <- hh.sizes[i]}
 }
}

HHsizes <- t(matrix(as.numeric(rep(HHsizes,d),1),n,d)); colnames(HHsizes) <- sid

HH8gps <- HHsizes>=8; colnames(HH8gps) <- sid

#Extract individual ID: sid and serial
SID <- t(as.matrix(as.numeric(substr(sid,4,7))))
Serial <- t(as.matrix(as.numeric(substr(sid,6,7))))

#Susceptibility matrix
sus <- matrix(1, d,n ) #dummy matrix

#Shedding matrix
shedding <-(rsv[,]==1)*1 ; colnames(shedding) <- sid

#Infection // a day before the onset
s.diff <- diff(shedding); infection <-rbind(s.diff, 0) ;
infection <- infection>0 ; colnames(infection) <- sid

#Recovery
onset <-rbind(0,s.diff) ; rm(ls=s.diff)
recover <- onset<0 ; colnames(recover) <- sid

#Onset ... allow multiple infections (reinfections)
onset <- onset>0; colnames(onset) <- sid

#Latency periods
pdf <- c(0, 0, 4, 4, 3, 1)/12
latency <- onset*1
```

```
for (c in 1:n) {
 for (r in 1:d) {
   if (onset[r,c]==1) {latency[r,c] <- pdf[1] }
   if (onset[r,c]==1) {latency[r-1,c] <- pdf[2] }
   if (onset[r,c]==1) {latency[r-2,c] <- pdf[3] }
   if (onset[r,c]==1) {latency[r-3,c] <- pdf[4] }
   if (onset[r,c]==1) {latency[r-4,c] <- pdf[5] }
   if (onset[r,c]==1) {latency[r-5,c] <- pdf[6] }
   }
}
onset.latency <- latency>0

#Post infection matrix
postInf <- matrix(0,d,n)
osre <- onset+recover; osre.t <- t(osre) ; nosre.t <- osre.t
for(i in 2:ncol(nosre.t)){
 nosre.t[,i] <- apply(osre.t[,1:i],1,sum)
}
nosre <- t(nosre.t)
for (c in 1:n) {
 for (r in 1:d) {
   if (nosre[r, c]>1) { postInf[r, c] <-1 }
 }
}

postInf <- postInf==1 ; colnames(postInf) <- sid
rm(ls=nosre.t); rm(ls=osre); rm(ls=osre.t); rm(ls=nosre)

# Status denotes when persons were away (0), or within household(1)
status <- as.matrix(rsv)
status <- status!=-1; status <- status*1

#At risk when on observation and not infected
atRisk <- status==1 & shedding!=1 ; colnames(atRisk) <- sid

# Present
present <- onset | atRisk

# Absent
absent <- !present

#Contact matrix // infants and others
Infant <- as.matrix(agegp0_1*1)
Infant.t <- t(Infant)

require(Matrix)
for (i in 1:47) {
 assign(names[i], matrix(1,hh.sizes[i],hh.sizes[i]))
}
```

```
Rlns <-
as.matrix(bdiag(hh_01,hh_03,hh_05,hh_06,hh_10,hh_11,hh_12,hh_14,hh_16,hh_17,
        hh_18,hh_19,hh_20,hh_21,hh_22,hh_23,hh_25,hh_26,hh_27,hh_28,
        hh_29,hh_30,hh_31,hh_33,hh_34,hh_35,hh_36,hh_38,hh_39,hh_40,
        hh_41,hh_42,hh_43,hh_44,hh_45,hh_46,hh_47,hh_48,hh_49,hh_51,
        hh_53,hh_54,hh_55,hh_56,hh_57,hh_58,hh_60))

Rlnship <- Rlns
colnames(Rlns) <- sid

for (r in 1:n ) {
 for (c in 1:n) {
  if (Rlns[r,c]==1 & Infant[c]==1 & Infant.t[r]==1) {Rlns[r,c] <- 1}
  if (Rlns[r,c]==1 & Infant[c]==0 & Infant.t[r]==1) {Rlns[r,c] <-2}
  if (Rlns[r,c]==1 & Infant[c]==1 & Infant.t[r]==FALSE) {Rlns[r,c] <-3}
  if (Rlns[r,c]==1 & Infant[c]==0 & Infant.t[r]==FALSE) {Rlns[r,c] <-4}
  if (r==c) {Rlns[r,c] <- 0 }
 }
}

#Relations status; [123][456]
rlns123gps <- Rlnship
for (r in 1:n ) {
 for (c in 1:n) {
  if (Rlnship[r,c]==1 & rlnship[c]>=1 & rlnship[c]<=3 & # child -> child
            rlnship.t[r]>=1 & rlnship.t[r]<=3)  {rlns123gps[r,c] <- 1}
    if (Rlnship[r,c]==1 & rlnship[c]>=4 & rlnship[c]<=6 &  # other -> child
            rlnship.t[r]>=1 & rlnship.t[r]<=3)  {rlns123gps[r,c] <- 2}
  if (Rlnship[r,c]==1 & rlnship[c]>=1 & rlnship[c]<=3 &  # child -> other
            rlnship.t[r]>=4 & rlnship.t[r]<=6)  {rlns123gps[r,c] <- 3}
  if (Rlnship[r,c]==1 & rlnship[c]>=4 & rlnship[c]<=6 &  # other -> other
            rlnship.t[r]>=4 & rlnship.t[r]<=6)  {rlns123gps[r,c] <- 4}
  if (r==c) {Rlns[r,c] <- 0 }
 }
}
colnames(rlns123gps) <- sid

#Relations nuclear vs. non-nuclear family ; [1245][36]
rlnsFam <- Rlnship
for (r in 1:n ) {
 for (c in 1:n) {
  if (Rlnship[r,c]==1 & rlnship[c]>=1 & rlnship[c]<=5 & rlnship[c]!=3 & #
                                                    family -> family
      rlnship.t[r]>=1 & rlnship.t[r]<=5 & rlnship.t[r]!=3)  {rlnsFam[r,c] <- 1}
  if (Rlnship[r,c]==1 & (rlnship[c]==3 | rlnship[c]==6) &  # other -> family
      rlnship.t[r]>=1 & rlnship.t[r]<=5 & rlnship.t[r]!=3)  {rlnsFam[r,c] <- 2}
  if (Rlnship[r,c]==1 & rlnship[c]>=1 & rlnship[c]<=5 & rlnship[c]!=3 &  #
                                                    family -> other
```

```
      (rlnship.t[r]==3 | rlnship.t[r]==6))  {rlnsFam[r,c] <- 3}
   if (Rlnship[r,c]==1 & (rlnship[c]==3 | rlnship[c]==6) &  # other -> other
      (rlnship.t[r]==3 | rlnship.t[r]==6))  {rlnsFam[r,c] <- 4}
   if (r==c) {Rlns[r,c] <- 0 }
  }
}
colnames(rlnsFam) <- sid

#Relations status; infant, children and adults [1][23][456]
rlns3gps <- Rlnship
for (r in 1:n ) {
  for (c in 1:n) {
    if (Rlnship[r,c]==1 & (rlnship[c]==1 | Infant[c]==1) &
                  (rlnship.t[r]==1 | Infant.t[r]==1)) {rlns3gps[r,c] = 1}
    if (Rlnship[r,c]==1 & rlnship[c]>=2 & rlnship[c]<=3 &  # child -> infant
                   rlnship.t[r]==1) {rlns3gps[r,c] <- 2}
    if (Rlnship[r,c]==1 & rlnship[c]>=4 & rlnship[c]<=6 &  # other -> infant
                  rlnship.t[r]==1) {rlns3gps[r,c] <- 3}
    if (Rlnship[r,c]==1 & rlnship[c]==1 & # infant -> child
                  rlnship.t[r]>=2 & rlnship.t[r]<=3) {rlns3gps[r,c] <- 4}
    if (Rlnship[r,c]==1 & rlnship[c]>=2 & rlnship[c]<=3 &  # child -> child
                  rlnship.t[r]>=2 & rlnship.t[r]<=3) {rlns3gps[r,c] <- 5}
    if (Rlnship[r,c]==1 & rlnship[c]>=4 & rlnship[c]<=6 &  # other -> child
                  rlnship.t[r]>=2 & rlnship.t[r]<=3) {rlns3gps[r,c] <- 6}
    if (Rlnship[r,c]==1 & rlnship[c]==1 & # infant -> other
                  rlnship.t[r]>=4 & rlnship.t[r]<=6) {rlns3gps[r,c] <- 7}
    if (Rlnship[r,c]==1 & rlnship[c]>=2 & rlnship[c]<=3 &  # child -> other
                  rlnship.t[r]>=4 & rlnship.t[r]<=6) {rlns3gps[r,c] <- 8}
    if (Rlnship[r,c]==1 & rlnship[c]>=4 & rlnship[c]<=6 &  # other -> other
                  rlnship.t[r]>=4 & rlnship.t[r]<=6) {rlns3gps[r,c] <- 9}
    if (r==c) {Rlns[r,c] <- 0 }
  }
}
colnames(rlns3gps) <- sid

# Exposure to virus in the household
exposure <- matrix(0,d,n)
for (c in 1:n) {
  ind1 <- Rlns[c,]==1
  ind2 <- Rlns[c,]==2
  ind3 <- Rlns[c,]==3
  ind4 <- Rlns[c,]==4
  for (r in 1:d) {
    exposure[r,c] <- (sum(shedding[r,ind1]))+(sum(shedding[r,ind2]))+
      (sum(shedding[r,ind3]))+sum(shedding[r,ind4])
  }
}
colnames(exposure) <- sid
```

```
#Community infection varying; shedding data summed across all individuals
virus <- matrix(apply(shedding, 1, sum))

quartz()
plot (1:d,virus,
    type='o',
    xlim=c(1,200),
    xlab='Days',
    ylab='No. of individuals infected',
   col="blue")

# lines(sampledates, virus,
#      col='blue')

#Function for fitting time-dependent community infection rate
comInf <- matrix(0,dim(virus))

#Time dependent exposure from wider community
#Values for the function from Matlab by GFMedley
shed.fit <- matrix(0,dim(shedding))
a1 <- 9.387 ; b1 <- 141.3 ; c1 <- 14.77 ; a2 <- 30.5 ; b2 <- 108.1; c2 <- 35.78
for (r in 1:d) {
 shed.fit[r] <-a1*exp(-((r-b1)/c1)^2) + a2*exp(-((r-b2)/c2)^2)
}

lines(1:d, shed.fit,
    col='red',
    lwd=4)

#Insert a legend
legend(150, 35,c('Observed','Fitted'),lty=1,col=c('blue','red'))

shedDen <- shed.fit/shed.fit[which.max(shed.fit)]
```

****

III. Script Defining The Functions (funDefs.R)

```
####**********LOGLIKELIHOOD FUNCTION**********#####
#********** cal.nLL.... START***************
#Age, HH, 3 contact groups

#sus - the structures for susceptibility.
#sus.A are the age groups set for indexing into age parameters. sus.B are the post
infection groups set for indexing

cal.nLL <- function (pars, sus, d, n, atRisk, shedDen,  latency,
                onset.latency, conParIndex,HH8gps) {
 logLike <- matrix(0, d, n)
```

```
#Parameters to estimate
 pC <- exp(pars[1]) #community
 pHH1 <- exp(pars[2]) #beta
 pHH2 <- exp(pars[16]) #HH

#  Age effect
 agePars <- c ( 1, exp(pars[4:7]))
 susA <- matrix ( agePars [ sus.A ], nrow=d, ncol=n, byrow=TRUE )

# Make the PI matrix and combine the susceptibility effects
 piPars <- c ( 1, exp(pars[3]))
 sus <- matrix ( piPars [ sus.B ]*susA, nrow=d, ncol=n )

# Contact structure
 conPars = c ( 0, 1, exp(pars[8:15]))
 exposure <- matrix(0,d,n)
 for (i in 1:n) { #for each individual
  ee <- conPars[ conParIndex[ i, ] ] #exposure to each individual
  exposure[,i] <- shedding %*% ee  #matrix multiplication - probably faster if
                                   #shedding was a sparse array
 }

# Force of infection
 lambdaC <- matrix (pC*shedDen,d,n) # varying community transmission rate over
                                    # time
lambda <- lambdaC + (pHH1*exposure*(1*(!HH8gps))) +  (pHH2*
       exposure*(HH8gps*1))   # community+HH infection rate

# Calculate negative log likelihood
 prInf <- log(1-exp(-lambda)) + log(sus)
 prNInf<- log(1-sus*(1-exp(-lambda)))
 logLike[atRisk] <- prNInf[atRisk]
 logLike[onset.latency] <- latency[onset.latency]*prInf[onset.latency]
 return (-sum(logLike))
}
#********** cal.nLL.....END**********************************#####

#***********cal.contrib....START*********************************######
#Function to calculate the contributions to observed infection –
# uses the same logic as the LL calculation
#Must supply the rlnship vector
#Works for a particular onset - defined as INDIV, DAY
cal.contrib <- function (pars, d, n, atRisk, shedDen,  latency,
                onset.latency, postInf, conParIndex,HH8gps, rlnship, onsets) {
 # pars - the model parameters (optimised already)
 # conParIndex - the index to the contact parameters+1;
#i.e. "2" means the value is the 1st parameter; diagonals 1 t

#Parameters
```

```
numOnsets = dim ( onsets )[1]
 p0 <- pars[1]
 p1 <- pars[2]
 p3 <- pars[3]
 p4 <- pars[4]
 p5 <- pars[5]
 p6 <- pars[6]
 p7 <- pars[7]
 cp = c(0,1,exp(pars[8:15])) #make an array of the contact parameters
 p8 <- pars[16]

 nDays = 7  #number of days latency considered over
 contrib <- matrix( 0, 205, 7 ) #the matrix of contributions, final column is the
                                # community contribution

# Calculate the force of infection (FOI) from the community
 lambdaC <- matrix(exp(p0)*shedDen,d,1) # varying community transmission rate
                                        # over time
   for (ii in 1:dim(onsets)[1]) { # for each individual onset
INDIV = onsets[ii,"INDIV"]
        DAY = onsets[ii,"DAY"]
     dayRange <- (DAY-6):DAY
     commCon <- sum(lambdaC[dayRange] * latency[dayRange,INDIV])

   # Generate the exposure and FOI from the HHs
   contacts <- which ( conParIndex[INDIV,]>1) #people who might have infected
                                              # this person
   nContacts <- length ( contacts )
   RR <- matrix ( conParIndex[INDIV,contacts], nrow=nDays, ncol=nContacts,
byrow=TRUE )  #the section of the contact data
   exposure <- matrix ( cp[RR], nDays, nContacts )
   HH <- HH8gps [ DAY, INDIV ]  #the HH grouping for the HH of this individual
   lambdaH <- (exp(p1)*exposure)*(!HH) + (exp(p8)*exposure)*HH
   hhCon <-lambdaH * latency[ dayRange, INDIV ] * shedding [dayRange, contacts]
   sumCon <- sum(hhCon)+commCon
   hhCon <- colSums ( hhCon ) / sumCon
   commCon <- commCon / sumCon

        for ( jj in 1:nContacts ) {
    contrib[ ii, rlnship[ contacts[ jj ]]] <- contrib [ ii, rlnship[contacts[ jj ]]] + hhCon[
jj ] }
   contrib[ ii, 7 ] <- commCon
    }
   return (contrib)
 }
 #**********************cal.contrib ....END*****************#######

#####################          END        ############################
```